

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 1, 2006, 22:34:01 ; Search time 42 Seconds
(without alignments)
1278.308 Million cell updates/sec

Title: US-09-189-415D-11
Perfect score: 558
Sequence: 1 MFIGNLGHNPVNNNSIPAP.....SNSAVNTSNNPAPGSHRTV 558

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 283416 seqs, 96216763 residues

Word size : 8

Total number of hits satisfying chosen parameters: 210

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	558	100.0	558	2	A98199
2	558	100.0	558	2	B86045
3	13	2.3	166	2	C90029
4	12	2.2	139	2	D86417
5	12	2.2	458	2	T31631
6	12	2.2	524	2	S33640
7	12	2.2	560	2	T32661
8	12	2.2	569	2	S47277
9	12	2.2	802	2	A36910
10	12	2.2	1023	2	S12519
11	12	2.2	3712	2	S18253
12	12	2.2	4377	2	A55575
13	11	2.0	67	2	B56888
14	11	2.0	108	2	T26880
15	11	2.0	164	2	T26861
16	11	2.0	183	2	S05358
17	11	2.0	208	2	T46896
18	11	2.0	217	2	S01358
19	11	2.0	234	2	T26560
20	11	2.0	245	2	T26868
21	11	2.0	274	2	A45632
22	11	2.0	278	2	S39310
23	11	2.0	284	2	T22023
24	11	2.0	304	2	T15922
25	11	2.0	327	2	S20074
26	11	2.0	341	2	T32949
27	11	2.0	342	2	T29557
28	11	2.0	373	2	T29596
29	11	2.0	385	2	JC7783

30	11	2.0	415	2	T32467	hypothetical prote
31	11	2.0	468	2	A55476	protein kinase (EC
32	11	2.0	477	2	A54643	nemo, form I - fru
33	11	2.0	484	2	S58868	G protein-coupled
34	11	2.0	512	2	T02498	probable WRKY-type
35	11	2.0	516	2	S19252	1-aminocyclopropan
36	11	2.0	517	2	T20658	probable zinc meta
37	11	2.0	518	2	S31442	1-aminocyclopropan
38	11	2.0	519	2	T23739	hypothetical prote
39	11	2.0	525	2	A35596	hypothetical prote
40	11	2.0	526	2	A56573	nuclear pore compl
41	11	2.0	530	2	T32812	hypothetical prote
42	11	2.0	551	2	S18408	alkaline phosphata
43	11	2.0	559	2	B36307	alkaline phosphata
44	11	2.0	651	2	T21175	hypothetical prote
45	11	2.0	680	2	T19939	hypothetical prote
46	11	2.0	681	2	T23454	hypothetical prote
47	11	2.0	698	2	A54796	regulatory protein
48	11	2.0	732	2	T25937	hypothetical prote
49	11	2.0	770	2	T22808	hypothetical prote
50	11	2.0	781	2	S51592	Xyln precursor - R
51	11	2.0	825	2	T29634	hypothetical prote
52	11	2.0	831	2	T08611	hypothetical prote
53	11	2.0	889	2	A35679	rep protein - slim
54	11	2.0	975	2	T08606	protein phosphatas
55	11	2.0	1002	2	T30546	major surface glyc
56	11	2.0	1076	2	JC2217	major surface glyc
57	11	2.0	1083	2	JC2300	cell surface glyco
58	11	2.0	1089	2	T14576	nosa protein - sl1
59	11	2.0	1093	2	T18275	1-phosphatidylinos
60	11	2.0	1099	2	T18257	phospholipase C -
61	11	2.0	1271	2	D64237	hypothetical prote
62	11	2.0	1282	2	UB0120	glycoprotein A - m
63	11	2.0	1402	2	T17456	cell surface prote
64	11	2.0	1635	2	T14075	cell surface prote
65	11	2.0	1671	2	S71628	senory transducti
66	11	2.0	1737	2	A59235	unconventional myo
67	11	2.0	1832	2	T31113	mucin-like glycopr
68	11	2.0	1858	2	T18273	1-phosphatidylinos
69	11	1.8	127	2	T51538	nitrilase associat
70	10	1.8	187	2	T49491	hypothetical prote
71	10	1.8	195	2	T19617	hypothetical prote
72	10	1.8	263	2	S01360	salivary glue prot
73	10	1.8	307	1	GSPP3	salivary glue prot
74	10	1.8	388	2	T16861	hypothetical prote
75	10	1.8	390	2	T49619	hypothetical prote
76	10	1.8	393	2	B86189	protein T25N20.9 l
77	10	1.8	395	2	T45599	hypothetical prote
78	10	1.8	435	2	T25350	hypothetical prote
79	10	1.8	572	2	T16865	hypothetical prote
80	10	1.8	577	2	G89430	hypothetical prote
81	10	1.8	648	1	JQ1150	protein K02E2.3 (l
82	10	1.8	649	2	T24505	protein kinase (BC
83	10	1.8	662	2	A45155	hypothetical prote
84	10	1.8	788	2	S05661	mucin F1M-C.1 - Af
85	10	1.8	816	2	CG9493	muscarinic acetylch
86	10	1.8	947	2	T08605	hypothetical prote
87	10	1.8	1008	2	T30544	hypothetical prote
88	10	1.8	1014	2	T18759	major surface glyco
89	10	1.8	1017	2	T30542	hypothetical prote
90	10	1.8	1022	2	T30543	major surface glyco
91	10	1.8	1030	2	T18374	major surface glyco
92	10	1.8	1390	2	T14004	B-cell receptor pr
93	10	1.8	1513	2	T23681	trif protein - all
94	10	1.8	1570	2	T18272	hypothetical prote
95	9	1.6	124	2	T48833	1-phosphatidylinos
96	9	1.6	202	2	F86755	hypothetical prote
97	9	1.6	213	2	T22865	prophage p12 prote
98	9	1.6	224	2	G86148	hypothetical prote
99	9	1.6	232	2	A60095	TfN6.14 protein -
100	9	1.6	492	2	A41907	lateral glue protei
101	9	1.6	500	1	EPPE	methyl-Cpg-binding
102	9	1.6	622	2	T51223	zip protein precur


```

Db      121  VQGRNGETSVLSDGEYARLQSIDPEGKDKFVTGRRGAGAMVTVAADITFARGL 180
Qy      181  ELLEPGTGSKAGSGKVGELRESNSGAEENTETSTSTSLRSDPKLMLALGTVA 240
      181  ELLEPGTGSKAGSGKVGELRESNSGAEENTETSTSTSLRSDPKLMLALGTVA 240
Qy      241  GLIGLATGIVQALALTPEDSPPTTDDPAASAATFATRDQLTKEAFQNPDMQKNIDE 300
      241  GLIGLATGIVQALALTPEDSPPTTDDPAASAATFATRDQLTKEAFQNPDMQKNIDE 300
Db      301  LGNAIPSGVLKDDVANNIEQAKAAGEAKQOAIENNAQOKKYDEQAKROBELKXSSG 360
      301  LGNAIPSGVLKDDVANNIEQAKAAGEAKQOAIENNAQOKKYDEQAKROBELKXSSG 360
Qy      361  AGGSLGALLGGGIGVAVTAALHRKNQPVQOTTTTTTTTTTTSARTVENKPPANTPAOG 420
      361  AGGSLGALLGGGIGVAVTAALHRKNQPVQOTTTTTTTTTTTSARTVENKPPANTPAOG 420
Db      421  NVDTPESEDTEMRSSMASTSTFPDTSIGTVQNPYADVKTSLHDSQVPTSNSVTQ 480
      421  NVDTPESEDTEMRSSMASTSTFPDTSIGTVQNPYADVKTSLHDSQVPTSNSVTQ 480
Qy      481  NMGNDSVYVSTIQHPPRDITDNGARLGNPSAGISTVARLALSGLRHDMGGLTGGSN 540
      481  NMGNDSVYVSTIQHPPRDITDNGARLGNPSAGISTVARLALSGLRHDMGGLTGGSN 540
Db      541  SAVNTSNNPPAPGSHRFV 558
      541  SAVNTSNNPPAPGSHRFV 558

```

RESULT 2

```

E86045
Probable translocated intilin receptor protein tir [imported] - Escherichia coli (strain
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C:Accession: E86045
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grobbeck, E.J.; Davis, N.W.; Llim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; UID:21074935; PMID:11206551
A:Accession: E86045
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-558 <STO>
A:Cross-references: UNIPROT:Q9R396; UNIPARC:UPI000000D00CA; GB:AE005174; NID:G12518449; F
A:Experimental source: strain O157:H7, eubacterial EDL933
C:Genetics:
A:Gene: tir

```

```

Query Match      100.0%; Score 558; DB 2; Length 558;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 558; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1  MPKGNIGHNPNVNNISIPPAPELPSQTDGAGGQGLINSTGPGSLALFTPPVNSMADSGD 60
      1  MPKGNIGHNPNVNNISIPPAPELPSQTDGAGGQGLINSTGPGSLALFTPPVNSMADSGD 60
Db      1  MPKGNIGHNPNVNNISIPPAPELPSQTDGAGGQGLINSTGPGSLALFTPPVNSMADSGD 60
Qy      61  NPASDVPGLPVNPMLAASEITLNDGEVYLHDHGPLDITNROISSVFVEVTEDEDKHTA 120
      61  NPASDVPGLPVNPMLAASEITLNDGEVYLHDHGPLDITNROISSVFVEVTEDEDKHTA 120
Db      61  NPASDVPGLPVNPMLAASEITLNDGEVYLHDHGPLDITNROISSVFVEVTEDEDKHTA 120
Qy      121  VQGRNGETSVLSDGEYARLQSIDPEGKDKFVTGRRGAGAMVTVAADITFARGL 180
      121  VQGRNGETSVLSDGEYARLQSIDPEGKDKFVTGRRGAGAMVTVAADITFARGL 180
Db      121  VQGRNGETSVLSDGEYARLQSIDPEGKDKFVTGRRGAGAMVTVAADITFARGL 180
Qy      181  ELLEPGTGSKAGSGKVGELRESNSGAEENTETSTSTSLRSDPKLMLALGTVA 240
      181  ELLEPGTGSKAGSGKVGELRESNSGAEENTETSTSTSLRSDPKLMLALGTVA 240
Db      181  ELLEPGTGSKAGSGKVGELRESNSGAEENTETSTSTSLRSDPKLMLALGTVA 240
Qy      241  GLIGLATGIVQALALTPEDSPPTTDDPAASAATFATRDQLTKEAFQNPDMQKNIDE 300
      241  GLIGLATGIVQALALTPEDSPPTTDDPAASAATFATRDQLTKEAFQNPDMQKNIDE 300

```

```

Db      241  GLIGLATGIVQALALTPEDSPPTTDDPAASAATFATRDQLTKEAFQNPDMQKNIDE 300
Qy      301  LGNAIPSGVLKDDVANNIEQAKAAGEAKQOAIENNAQOKKYDEQAKROBELKXSSG 360
      301  LGNAIPSGVLKDDVANNIEQAKAAGEAKQOAIENNAQOKKYDEQAKROBELKXSSG 360
Db      301  LGNAIPSGVLKDDVANNIEQAKAAGEAKQOAIENNAQOKKYDEQAKROBELKXSSG 360
Qy      361  AGGSLGALLGGGIGVAVTAALHRKNQPVQOTTTTTTTTTTTSARTVENKPPANTPAOG 420
      361  AGGSLGALLGGGIGVAVTAALHRKNQPVQOTTTTTTTTTTTSARTVENKPPANTPAOG 420
Db      421  NVDTPESEDTEMRSSMASTSTFPDTSIGTVQNPYADVKTSLHDSQVPTSNSVTQ 480
      421  NVDTPESEDTEMRSSMASTSTFPDTSIGTVQNPYADVKTSLHDSQVPTSNSVTQ 480
Qy      481  NMGNDSVYVSTIQHPPRDITDNGARLGNPSAGISTVARLALSGLRHDMGGLTGGSN 540
      481  NMGNDSVYVSTIQHPPRDITDNGARLGNPSAGISTVARLALSGLRHDMGGLTGGSN 540
Db      541  SAVNTSNNPPAPGSHRFV 558
      541  SAVNTSNNPPAPGSHRFV 558

```

RESULT 3

```

C90029
Hypothetical protein SA2097 [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C:Accession: C90029
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Chii, L.; Ogu
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A89758; UID:21311952; PMID:11418146
A:Accession: C90029
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-166 <RUR>
A:Cross-references: UNIPROT:Q99RM9; UNIPARC:UPI000005434E; GB:BA000018; PID:G13702104;
A:Experimental source: strain N315
C:Genetics:
A:Gene: SA2097

```

```

Query Match      2.3%; Score 13; DB 2; Length 166;
Best Local Similarity 100.0%; Pred. No. 0.00015;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      392  QTTTTTTTTTTTTT 404
      43  QTTTTTTTTTTTTT 55
Db      43  QTTTTTTTTTTTTT 55

```

RESULT 4

```

D86417
Probable auxin-induced protein, 50455-50036 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Dec-2004
C:Accession: D86417
R:Rheologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.
ansen, N.F.; Hughes, B.; Hultzer, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Klm, C
A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luos, J.S.; Maiti, R.; Marshall,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Frazer, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; UID:21016719; PMID:11130712
A:Accession: D86417
A:Status: preliminary

```

A:Molecule type: DNA

A:Residues: 1-139 <STO>

A:Cross-references: UNIPROT:Q9CTQ5; UNIPARC:UPI000000AA867; GB:AE005172; NID:g10092232; F

C:Genetics:

A:Map position: 1

C:Superfamily: auxin-induced protein 10A

Query Match 2.2%; Score 12; DB 2; Length 139;

Best Local Similarity 100.0%; Pred. No. 0.0012;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 393 TTTT TTTT TTTT TTTT 404

DB 22 TTTT TTTT TTTT TTTT 33

RESULT 5

T31631

hypothetical protein Y57A10A.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004

C:Accession: T31631

R:Smyle, R.

submitted to the EMBL Data Library, September 1999

A:Reference number: 221048

A:Accession: T31631

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-458 <WIL>

A:Cross-references: UNIPROT:Q9NAB3; UNIPARC:UPI0000082499; EMBL:AL117195; NID:e1549729;

A:Experimental source: clone Y57A10A

C:Genetics:

A:Gene: CESP:Y57A10A.1

A:introns: 8/3; 54/3; 112/3; 151/1

Query Match 2.2%; Score 12; DB 2; Length 458;

Best Local Similarity 100.0%; Pred. No. 0.0035;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 393 TTTT TTTT TTTT TTTT 404

DB 135 TTTT TTTT TTTT TTTT 146

RESULT 6

S33640

homeotic protein smox-2, engrailed-like - fluke (Schistosoma mansoni)

C:Species: Schistosoma mansoni

C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 05-Oct-2004

C:Accession: S33640; S27841

R:Webster, P.J.; Mansour, T.E.

Mech. Dev. 38, 25-32, 1992

A:Title: Conserved classes of homeodomains in Schistosoma mansoni, an early bilateral me

A:Reference number: S33640; MUID:92399260; PMID:1356008

A:Accession: S33640

A:Molecule type: mRNA

A:Residues: 1-524 <MEB>

A:Cross-references: UNIPROT:Q26601; UNIPARC:UPI000012CA06; EMBL:S44191; EMBL:M65305; NID

C:Genetics:

A:Gene: smox-2

C:Keywords: DNA binding; homeobox; nucleus; transcription regulation

F:424-480/Domain: homeobox homology <HOX>

Query Match 2.2%; Score 12; DB 2; Length 524;

Best Local Similarity 100.0%; Pred. No. 0.004;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 393 TTTT TTTT TTTT TTTT 404

DB 105 TTTT TTTT TTTT TTTT 116

RESULT 7

T32661

hypothetical protein K11D12.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999

C:Accession: T32661

R:Henkhaus, J.; Wohldmann, P.; Gillam, B.

submitted to the EMBL Data Library, December 1997

A:Description: The sequence of C. elegans coamid K11D12.

A:Reference number: 221207

A:Accession: T32661

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-560 <HEN>

A:Cross-references: UNIPARC:UPI000017BABA; EMBL:AF039047; PIDN:AA894223.1; GSPDB:GN00023

C:Genetics:

A:Gene: CESP:K11D12.1

A:Map position: 5

A:introns: 5/3; 48/3; 90/3; 127/3; 149/3; 190/1; 207/1; 233/3; 264/1; 480/1

Query Match 2.2%; Score 12; DB 2; Length 560;

Best Local Similarity 100.0%; Pred. No. 0.0042;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 393 TTTT TTTT TTTT TTTT 404

DB 348 TTTT TTTT TTTT TTTT 359

RESULT 8

S47277

gp88 protein - murine cytomegalovirus

C:Species: murine cytomegalovirus; murine herpesvirus 1

C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004

C:Accession: S47277

R:Thiele, R.; Lucin, P.; Schneider, K.; Koszowski, U.

submitted to the EMBL Data Library, February 1994

A:Reference number: S47277

A:Accession: S47277

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-569 <THA>

A:Cross-references: UNIPROT:Q8J183; UNIPARC:UPI00000EFEBF; EMBL:X77798; NID:g535195; PID

C:Superfamily: murine cytomegalovirus gp88 protein

Query Match 2.2%; Score 12; DB 2; Length 569;

Best Local Similarity 100.0%; Pred. No. 0.0043;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 392 OTTT TTTT TTTT 403

DB 472 OTTT TTTT TTTT 483

RESULT 9

A36910

xylanase, beta(1,3-1,4)-glucanase - Ruminococcus flavefaciens

C:Species: Ruminococcus flavefaciens

C>Date: 07-Apr-1994 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004

C:Accession: A36910

R:Plint, H.J.; Martin, J.; McPherson, C.A.; Daniel, A.S.; Zhang, J.X.

J. Bacteriol. 175, 2943-2951, 1993

A:Title: A bifunctional enzyme, with separate xylinase and beta(1,3-1,4)-glucanase domai

A:Reference number: A36910; MUID:9325938; PMID:8491715

A:Accession: A36910

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-802 <FLI>

A:Cross-references: UNIPROT:Q9S310; UNIPARC:UPI00000B74F4; GB:S61204; NID:g385910; PIDN:

A:Note: sequence extracted from NCBI backbone (NCBIN:131871, NCBI:131872)

F:42-239/Domain: endo-1,4-beta-xylinase homology <XYL>

F:259-401/Domain: Thermotoga xylinase A amino-terminal repeat homology <TXA>

Query Match 2.2% Score 12; DB 2; Length 802;
 Best Local Similarity 100.0%; Pred. No. 0.0058;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 393 TTTT TTTT TTTT TTTT 404
 |||||
 Db 533 TTTT TTTT TTTT TTTT 544

RESULT 10
 S12519
 glutactin - fruit fly (Drosophila melanogaster)
 C:Species: Drosophila melanogaster
 C:Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
 C:Accession: S12519
 R:Olson, P.F.; Fessler, L.I.; Nelson, R.E.; Sterne, R.E.; Campbell, A.G.; Fessler, J.H.
 EMBL J. 9, 1219-1227, 1990
 A>Title: Glutactin, a novel Drosophila basement membrane-related glycoprotein with sequ
 A:Reference number: S12519; MUID:90214632; PMID:2108664
 A:Accession: S12519
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1023 <OLS>
 A:Cross-references: UNIPROT:P33438; UNIPARC:UPI000012B819; EMBL:X53286; NID:g297084; PID
 C:Genetics:
 A:introns: 390/3

Query Match 2.2% Score 12; DB 2; Length 1023;
 Best Local Similarity 100.0%; Pred. No. 0.0072;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 393 TTTT TTTT TTTT TTTT 404
 |||||
 Db 605 TTTT TTTT TTTT TTTT 616

RESULT 11
 S18253
 laminin alpha-1 chain precursor - fruit fly (Drosophila melanogaster)
 C:Species: Drosophila melanogaster
 C:Date: 16-Sep-1992 #sequence_revision 24-Jul-1997 #text_change 09-Jul-2004
 C:Accession: S28399; S18253
 R:Kunche-Gulberg, M.; Garrison, K.; Mackrell, A.J.; Fessler, L.I.; Fessler, J.H.
 EMBL J. 11, 4519-4527, 1992
 A>Title: Laminin A chain: expression during Drosophila development and genomic sequence.
 A:Reference number: S28399; MUID:93049203; PMID:1425586
 A:Accession: S28399
 A:Status: preliminary
 A:Molecule type: nucleic acid
 A:Residues: 1-3712 <KUS>
 A:Cross-references: UNIPROT:Q00174; UNIPARC:UPI000004EE1C; GB:M96388; NID:g157799; PIDN:
 R:Garrison, K.; Mackrell, A.J.; Fessler, J.H.
 J. Biol. Chem. 266, 22899-22904, 1991
 A>Title: Drosophila laminin A chain sequence, interspecies comparison, and domain struct
 A:Reference number: S18253; MUID:92078147; PMID:1744083
 A:Accession: S18253
 A:Molecule type: mRNA
 A:Residues: 1762-3712 <GAR>
 A:Cross-references: UNIPARC:UPI000016BC67; EMBL:M75882; NID:g157797; PIDN:AAA28661.1; PI
 C:Genetics:
 A:Gene: FlyBase:lanA
 A:Cross-references: FlyBase:FBgn0002526
 C:Superfamily: laminin alpha-1 chain; laminin G repeat homology; laminin-type EGF-like H
 C:Keywords: basement membrane; cell binding; coiled coil; disulfide bond; extracellular
 F:273-330/Domain: laminin-type EGF-like homology <LEG>
 F:333-400/Domain: laminin-type EGF-like homology <LEG2>
 F:541-584/Domain: laminin-type EGF-like homology <LEG1>
 F:1776-2115/Domain: IIT <DOM3>
 F:1776-1806/Domain: laminin-type EGF-like homology #status atypical <LEI>
 F:1809-1856/Domain: laminin-type EGF-like homology <LE2>
 F:1859-1914/Domain: laminin-type EGF-like homology <LE3>
 F:1917-1967/Domain: laminin-type EGF-like homology <LE4>
 F:1970-2014/Domain: laminin-type EGF-like homology <LE5>

F:2017-2061/Domain: laminin-type EGF-like homology <LE6>
 F:2064-2109/Domain: laminin-type EGF-like homology <LE7>
 F:2116-2637/Domain: I/II, heparin repeats <DOM2>
 F:2698-3712/Domain: G <DOMG>
 F:2698-2863/Domain: repeat G1 <RG1>
 F:2864-3048/Domain: repeat G2 <RG2>
 F:3049-3223/Domain: repeat G3 <RG3>
 F:3079-3200/Domain: laminin G repeat homology <LG3>
 F:3334-3528/Domain: repeat G4 <RG4>
 F:3529-3712/Domain: repeat G5 <RG5>
 F:1847,1850,1943,2024,2196,2215,2267,2301,2323,2482,2524,2538,2569,2699,2720,2890,2938,

Query Match 2.2% Score 12; DB 2; Length 3712;
 Best Local Similarity 100.0%; Pred. No. 0.022;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 393 TTTT TTTT TTTT TTTT 404
 |||||
 Db 3273 TTTT TTTT TTTT TTTT 3284

RESULT 12
 A55575
 ankryrin 3, long splice form - human
 N:Alternate names: ankryrin G
 C:Species: Homo sapiens (man)
 C:Date: 03-Mar-1995 #sequence_revision 03-Mar-1995 #text_change 31-Dec-2004
 C:Accession: A55575
 R:Kordella, E.; Lambert, S.; Bennett, V.
 J. Biol. Chem. 270, 2352-2359, 1995
 A>Title: Ankryrin-G, A new ankryrin gene with neural-specific isoforms localized at the a
 A:Reference number: A55575; MUID:95138209; PMID:7836469
 A:Accession: A55575
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-4377 <KOR>
 A:Cross-references: UNIPROT:Q12955; UNIPARC:UPI000013C497; GB:U13616; NID:g608024; PIDN
 C:Genetics:
 A:Gene: GDB:ANK3
 A:Cross-references: GDB:424503; OMIM:600465
 A:Map position: 10q21-10q21
 C:Superfamily: EGF homology
 C:Keywords: alternative splicing; peripheral membrane protein
 F:73-105/Domain: ankryrin repeat homology <AN01>
 F:106-138/Domain: ankryrin repeat homology <AN02>
 F:139-171/Domain: ankryrin repeat homology <AN03>
 F:172-200/Domain: ankryrin repeat homology <AN04>
 F:201-233/Domain: ankryrin repeat homology <AN05>
 F:234-266/Domain: ankryrin repeat homology <AN06>
 F:267-299/Domain: ankryrin repeat homology <AN07>
 F:300-332/Domain: ankryrin repeat homology <AN08>
 F:333-365/Domain: ankryrin repeat homology <AN09>
 F:366-398/Domain: ankryrin repeat homology <AN10>
 F:399-431/Domain: ankryrin repeat homology <AN11>
 F:432-464/Domain: ankryrin repeat homology <AN12>
 F:465-497/Domain: ankryrin repeat homology <AN13>
 F:498-530/Domain: ankryrin repeat homology <AN14>
 F:531-563/Domain: ankryrin repeat homology <AN15>
 F:564-596/Domain: ankryrin repeat homology <AN16>
 F:597-629/Domain: ankryrin repeat homology <AN17>
 F:630-662/Domain: ankryrin repeat homology <AN18>
 F:663-695/Domain: ankryrin repeat homology <AN19>
 F:696-728/Domain: ankryrin repeat homology <AN20>
 F:729-761/Domain: ankryrin repeat homology <AN21>
 F:762-794/Domain: ankryrin repeat homology <AN22>
 F:795-827/Domain: ankryrin repeat homology <AN23>

Query Match 2.2% Score 12; DB 2; Length 4377;
 Best Local Similarity 100.0%; Pred. No. 0.026;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 393 TTTT TTTT TTTT TTTT 404
 |||||

Db 3968 TTTTTTTTTTTS 3979

RESULT 13

B56888

alkaline phosphatase (EC 3.1.3.1), intestinal type II - rat (fragment)

C:Species: Rattus norvegicus (Norway rat)

C:Date: 05-Jan-1996 #sequence_revision 05-Jan-1996 #text_change 16-Aug-2004

C:Accession: B56888

R:Engle, M.J.; Alpers, D.H.

Clin. Chem. 38, 2506-2509, 1992

A:Title: The two mRNAs encoding rat intestinal alkaline phosphatase represent two distinct

A:Reference number: A56888; MUID:93092310; PMID:1458592

A:Accession: B56888

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-67 <ENG>

A:Cross-references: UNIPARC:UPI0000175940

A:Experimental source: duodenal mucosa

A:Note: sequence extracted from NCBI backbone (NCBIN:121249, NCBIPI:121252)

C:Superfamily: Alkaline phosphatase

C:Keywords: intestine; membrane protein; phosphoric monoester hydrolase

Query Match 2.0%; Score 11; DB 2; Length 67;

Best Local Similarity 100.0%; Pred. No. 0.0061;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 393 TTTTTTTTTTTT 403

|||||

Db 27 TTTTTTTTTTTT 37

RESULT 14

T26880

hypochemical protein Y43F8C.9 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C:Accession: T26880

R:Ainscough, R.

submitted to the EMBL Data Library, October 1998

A:Reference number: Z20279

A:Accession: T26880

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-108 <WIL>

A:Cross-references: UNIPROT:O9XWNO; UNIPARC:UPI00000612AC; EMBL:AL032637; PIDN:CAA21621.

A:Experimental source: clone Y43F8C

C:Genetics:

A:Gene: CESP:Y43F8C.9

A:introns: 40/3

Query Match 2.0%; Score 11; DB 2; Length 108;

Best Local Similarity 100.0%; Pred. No. 0.0093;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 393 TTTTTTTTTTTT 403

|||||

Db 50 TTTTTTTTTTTT 60

RESULT 15

T26561

hypochemical protein Y24F12A.d - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 08-Sep-2000

C:Accession: T26561

R:Lenard, N.

submitted to the EMBL Data Library, September 1999

A:Reference number: Z20233

A:Accession: T26561

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-164 <WIL>

A:Cross-references: UNIPARC:UPI0000179900; EMBL:AL110480; PIDN:CAB54380.1; CESP:Y24F12A.

A:Experimental source: clone Y24F12A

C:Genetics:

A:Gene: CESP:Y24F12A.d

A:introns: 137/1

C:Superfamily: Caenorhabditis elegans hypothetical protein Y9D1A.2

Query Match 2.0%; Score 11; DB 2; Length 164;

Best Local Similarity 100.0%; Pred. No. 0.013;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 393 TTTTTTTTTTTT 403

|||||

Db 112 TTTTTTTTTTTT 122

Search completed: August 1, 2006, 22:38:56

Job time : 50 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 1, 2006, 21:37:50 : Search time 195 Seconds
(without alignments)
70.341 Million cell updates/sec

Title: US-09-189-415D-7

Perfect score: 165
Sequence: 1 PIGNLGNVNGNHLIPAPPLPQTGDAAR 30

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 2589679 seqs, 457216429 residues

Total number of hits satisfying chosen parameters: 2589679

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 100 summaries

Database :

A_Geneseq_8: *
1: geneseqp1980s: *
2: geneseqp1990s: *
3: geneseqp2000s: *
4: geneseqp2001s: *
5: geneseqp2002s: *
6: geneseqp2003s: *
7: geneseqp2004s: *
8: geneseqp2005s: *
9: geneseqp2006s: *
10: geneseqp2007s: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	165	100.0	30	2	AAV06213
2	165	100.0	549	2	AAV06220
3	123	74.5	558	7	ADCO0799
4	123	74.5	558	9	ABP08155
5	123	74.5	558	10	ABP08155
6	123	74.5	558	10	ABP08155
7	123	74.5	558	10	ABP08155
8	123	74.5	558	10	ABP08155
9	123	74.5	558	10	ABP08155
10	123	74.5	558	10	ABP08155
11	123	74.5	558	10	ABP08155
12	123	74.5	558	10	ABP08155
13	123	74.5	558	10	ABP08155
14	123	74.5	558	10	ABP08155
15	123	74.5	558	10	ABP08155
16	123	74.5	558	10	ABP08155
17	123	74.5	558	10	ABP08155
18	123	74.5	558	10	ABP08155
19	123	74.5	558	10	ABP08155
20	123	74.5	558	10	ABP08155
21	123	74.5	558	10	ABP08155
22	123	74.5	558	10	ABP08155
23	123	74.5	558	10	ABP08155

24	56	33.9	410	4	ABP96146	ABP96146 Putative
25	56	33.9	410	8	ADP43142	ADP43142 Bacterial
26	56	33.9	474	7	ABP06765	ABP06765 Klebsiella
27	56	33.9	662	9	ADP18205	ADP18205 E. grandis
28	55.5	33.6	301	8	ADP18205	ADP18205 E. grandis
29	55.5	33.6	627	10	ADP18205	ADP18205 E. grandis
30	55	33.3	319	8	ADP18205	ADP18205 E. grandis
31	55	33.3	598	6	ADP18205	ADP18205 E. grandis
32	55	33.3	631	7	ADP18205	ADP18205 E. grandis
33	55	33.3	899	7	ADP18205	ADP18205 E. grandis
34	54.5	33.0	283	4	ADP18205	ADP18205 E. grandis
35	54.5	33.0	618	7	ADP18205	ADP18205 E. grandis
36	54.5	33.0	207	7	ADP18205	ADP18205 E. grandis
37	54	32.7	226	7	ADP18205	ADP18205 E. grandis
38	54	32.7	279	7	ADP18205	ADP18205 E. grandis
39	54	32.7	279	9	ADP18205	ADP18205 E. grandis
40	54	32.7	319	8	ADP18205	ADP18205 E. grandis
41	54	32.7	367	8	ADP18205	ADP18205 E. grandis
42	54	32.7	454	8	ADP18205	ADP18205 E. grandis
43	54	32.7	473	4	ADP18205	ADP18205 E. grandis
44	54	32.7	623	7	ADP18205	ADP18205 E. grandis
45	54	32.7	1238	4	ADP18205	ADP18205 E. grandis
46	54	32.7	1965	6	ADP18205	ADP18205 E. grandis
47	54	32.7	1966	6	ADP18205	ADP18205 E. grandis
48	54	32.7	1966	6	ADP18205	ADP18205 E. grandis
49	54	32.7	1966	6	ADP18205	ADP18205 E. grandis
50	54	32.7	1966	6	ADP18205	ADP18205 E. grandis
51	54	32.7	1966	6	ADP18205	ADP18205 E. grandis
52	54	32.7	1966	6	ADP18205	ADP18205 E. grandis
53	54	32.7	1966	6	ADP18205	ADP18205 E. grandis
54	54	32.7	1966	6	ADP18205	ADP18205 E. grandis
55	54	32.7	1966	6	ADP18205	ADP18205 E. grandis
56	54	32.7	1966	6	ADP18205	ADP18205 E. grandis
57	54	32.7	1966	6	ADP18205	ADP18205 E. grandis
58	54	32.7	1966	6	ADP18205	ADP18205 E. grandis
59	54	32.7	1966	6	ADP18205	ADP18205 E. grandis
60	54	32.7	1966	6	ADP18205	ADP18205 E. grandis
61	54	32.7	1966	6	ADP18205	ADP18205 E. grandis
62	54	32.7	1966	6	ADP18205	ADP18205 E. grandis
63	54	32.7	1966	6	ADP18205	ADP18205 E. grandis
64	54	32.7	1966	6	ADP18205	ADP18205 E. grandis
65	54	32.7	1966	6	ADP18205	ADP18205 E. grandis
66	54	32.7	1966	6	ADP18205	ADP18205 E. grandis
67	54	32.7	1966	6	ADP18205	ADP18205 E. grandis
68	54	32.7	1966	6	ADP18205	ADP18205 E. grandis
69	54	32.7	1966	6	ADP18205	ADP18205 E. grandis
70	54	32.7	1966	6	ADP18205	ADP18205 E. grandis
71	54	32.7	1966	6	ADP18205	ADP18205 E. grandis
72	54	32.7	1966	6	ADP18205	ADP18205 E. grandis
73	54	32.7	1966	6	ADP18205	ADP18205 E. grandis
74	54	32.7	1966	6	ADP18205	ADP18205 E. grandis
75	54	32.7	1966	6	ADP18205	ADP18205 E. grandis
76	54	32.7	1966	6	ADP18205	ADP18205 E. grandis
77	54	32.7	1966	6	ADP18205	ADP18205 E. grandis
78	54	32.7	1966	6	ADP18205	ADP18205 E. grandis
79	54	32.7	1966	6	ADP18205	ADP18205 E. grandis
80	54	32.7	1966	6	ADP18205	ADP18205 E. grandis
81	54	32.7	1966	6	ADP18205	ADP18205 E. grandis
82	54	32.7	1966	6	ADP18205	ADP18205 E. grandis
83	54	32.7	1966	6	ADP18205	ADP18205 E. grandis
84	54	32.7	1966	6	ADP18205	ADP18205 E. grandis
85	54	32.7	1966	6	ADP18205	ADP18205 E. grandis
86	54	32.7	1966	6	ADP18205	ADP18205 E. grandis
87	54	32.7	1966	6	ADP18205	ADP18205 E. grandis
88	54	32.7	1966	6	ADP18205	ADP18205 E. grandis
89	54	32.7	1966	6	ADP18205	ADP18205 E. grandis
90	54	32.7	1966	6	ADP18205	ADP18205 E. grandis
91	54	32.7	1966	6	ADP18205	ADP18205 E. grandis
92	54	32.7	1966	6	ADP18205	ADP18205 E. grandis
93	54	32.7	1966	6	ADP18205	ADP18205 E. grandis
94	54	32.7	1966	6	ADP18205	ADP18205 E. grandis
95	54	32.7	1966	6	ADP18205	ADP18205 E. grandis
96	54	32.7	1966	6	ADP18205	ADP18205 E. grandis

XX ADCC00799;
 AC 04-DEC-2003 (first entry)
 DT
 XX Enterohaemorrhagic E. coli O157:H7-specific protein SEQ ID NO: 844.
 XX
 XX enterohaemorrhagic; anti-bacterial.
 XX
 OS Escherichia coli; O157:H7.
 XX
 XX JP2002355074-A.
 XX
 PD 10-DEC-2002.
 XX
 XX 24-JAN-2002; 2002JP-00015959.
 XX
 XX 24-JAN-2001; 2001JP-00112010.
 XX
 XX (UYTS-) UNIV TSUKUBA.
 XX
 XX WPI; 2003-451640/43.
 DR
 XX Enterohaemorrhagic Escherichia coli O157:H7-specific nucleic acid molecule
 PT and a polypeptide and its use, a polypeptide, a vector and a host cell.
 PT
 XX Claim 3; SEQ ID NO 844; 2067PP; Japanese.
 PS
 CC The invention relates to a novel enterohaemorrhagic Escherichia coli
 CC O157:H7-specific nucleic acid molecule. A polynucleotide of the invention
 CC has anti-bacterial activity. The polypeptide can be used in detection
 CC and/or treatment of O157:H7 infection. The nucleotide sequence of the
 CC genome of Enterohaemorrhagic E. coli O157:H7 was determined. The present
 CC sequence represents an E. coli O157:H7-specific polypeptide of the
 CC invention.
 CC
 XX Sequence 558 AA;
 SQ

Query Match 74.5%; Score 123; DB 7; Length 558;
 Best Local Similarity 82.1%; Pred. No. 7.3e-07;
 Matches 23; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 PIGNLGNNVNGNHLIPAPPLPSQTDGA 28
 |||||:|:|:|||||||
 Db 2 PIGNLGHNPNVNSIPAPPLPSQTDGA 29

RESULT 4

ID AEB91310 standard; protein; 558 AA.
 AC
 XX AEB91310;
 XX

DT 20-OCT-2005 (first entry)
 XX

DE Microbial pathogen adhesin protein sequence, SEQ ID NO:20.
 XX

XX algorithm; adhesin; pharmaceutical; vaccine; drug screening;
 KW bordetella pertussis infection; antibacterial; pneumonia;
 XX antiinflammatory; respiratory-gen.; gastric ulcer; antidiabetic;
 KW gastroenteric-gen.; urinary tract infection; antimicrobial; uropathic.
 XX
 OS Escherichia coli.
 XX
 XX WO2005076010-A2.
 XX
 XX 18-AUG-2005.
 PD
 XX 07-FEB-2005; 2005WO-IN000037.
 XX
 XX 06-FEB-2004; 2004IN-DE000173.
 PR 20-JUL-2004; 2004US-0589227P.
 XX

PA (COUL) COUNCIL SCI & IND RES SOUTH AFRICA.
 XX
 XX Sachdeva G, Kumar K, Jain P, Brahmachari SK, Ramachandran S;
 XX
 DR WPI; 2005-597835/61.
 XX
 XX Computational method for identifying adhesin and adhesin like molecules,
 PT comprises computing sequence-based attributes of protein sequences using
 PT neural network software and training an artificial neural network.
 XX
 XX Claim 16; SEQ ID NO 20; 402pp; English.
 PS

CC The present invention relates to a computational method (M1) for
 CC identifying adhesin and adhesin-like proteins, by computing the sequence-
 CC based attributes of protein sequences using five attribute modules of a
 CC neural network software, training an artificial neural network (ANN) for
 CC each of the computed five attributes, and identifying the adhesin and
 CC adhesin-like proteins having probability of being an adhesin (Pad) as
 CC equal or greater than 0.51. Also claimed is a set of 274 annotated genes
 CC encoding adhesin and adhesin-like proteins, having 274 fully defined 162-
 CC 1151 base pair (SEQ ID NO: 385-658) sequences; a set of 105 hypothetical
 CC genes encoding adhesin and adhesin-like proteins, having 105 fully
 CC defined 306-15876 base pairs (SEQ ID NO: 659-763); a set of 279 annotated
 CC adhesin and adhesin-like proteins, having 279 fully defined 53-3716 base
 CC pair (SEQ ID NO: 1-279); a set of 105 hypothetical adhesin and adhesin-
 CC like proteins, having 105 fully defined 106-5291 base pair (SEQ ID NO:
 CC 280-384) sequences; and a fully connected multilayer feed forward ANN (1)
 CC based on (M1). (M1) is useful for identifying adhesin and adhesin-like
 CC proteins, of therapeutic potential, and identifying and short-listing
 CC proteins for further testing in development of new vaccine formulations
 CC to eliminate diseases caused by various pathogenic organisms. (M1) is
 CC useful for identifying putative adhesins that are important in drug
 CC discovery and preventing therapeutics for whooping cough, pneumonia,
 CC gastric ulcer and urinary tract infections. (M1) identifies adhesins from
 CC distantly related organisms, and from bacteria belonging to a wide
 CC phylogenetic spectrum. (M1) is capable of predicting adhesive nature of
 CC unique proteins. The present sequence is a microbial pathogen adhesin
 CC protein sequence.
 CC
 XX Sequence 558 AA;
 SQ

Query Match 74.5%; Score 123; DB 9; Length 558;
 Best Local Similarity 82.1%; Pred. No. 7.3e-07;
 Matches 23; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 PIGNLGNNVNGNHLIPAPPLPSQTDGA 28
 |||||:|:|:|||||||
 Db 2 PIGNLGHNPNVNSIPAPPLPSQTDGA 29

RESULT 5

ID AEE86220 standard; protein; 558 AA.
 AC
 XX AEE86220;
 XX

DT 23-FEB-2006 (first entry)
 XX

DE Escherichia coli translocated intimin receptor (Tir) protein.
 XX

XX diagnosis; therapeutic; screening; escherichia coli infection;
 KW antibacterial; infection; translocated intimin receptor.
 XX
 XX Escherichia coli.
 XX
 XX US2005287569-A1.
 XX
 XX 29-DEC-2005.
 PD
 XX 20-MAY-2005; 2005US-00134563.
 XX
 XX 20-MAY-2004; 2004US-0573600P.
 PR
 XX

PA (LEON/) LEONG J M.
 PA (CAMP/) CAMPBELLONE K G.
 XX
 XX Leong JM, Campellone KG;
 DR WPI: 2006-065745/07.
 DR N-PSDB; AEE86219.
 XX
 PT Novel purified polypeptide having six residues of EspF-U, and binding to
 PT neuronal Wiskott-Aldrich syndrome protein polypeptide, useful for
 PT identifying EspF-U activity modulating compound.
 PS Disclosure; SEQ ID NO 12; 62pp; English.
 XX
 CC The present invention relates to novel EspFu polypeptides and their
 CC corresponding polynucleotides. The EspFu polypeptides are EspF-like
 CC polypeptides encoded by genes of the cryptic prophage CP-933U of
 CC enterohemorrhagic Escherichia coli (EHEC) or enteropathogenic Escherichia
 CC coli (EPEC) that binds to a neuronal wiskott-aldrich syndrome protein (N-
 CC WASP) polypeptide or restores the actin pedestal formation activity of
 CC enteropathogenic E. coli (EPEC) strain K12. The invention further
 CC relates to a method of identifying a candidate compounds capable of
 CC binding to and/or modulating the activity of EspFu and compounds that
 CC inhibits protein-protein interactions between EspFu and EspFu-interacting
 CC proteins such as N-WASP, transducer of Cdc42-dependent actin assembly-1
 CC (Toca-1) and p21-activated kinase 1 (Pak1). EspFu polynucleotides are
 CC useful for diagnosing or detecting EHEC infection. EspFu antibody is
 CC useful for treating EHEC infection. The present sequence is the
 CC enterohemorrhagic Escherichia coli translocated intimin receptor (Tir)
 CC protein. This sequence is critical for the formation of actin pedestals
 CC in EHEC.
 SQ Sequence 558 AA;

Query Match 74.5%; Score 123; DB 10; Length 558;
 Best Local Similarity 82.1%; Pred. No. 7.3e-07;
 Matches 23; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 QY 1 PIGNLGNVNGNHLIPAPPLPSQTDGA 28
 |||||:|:|||||||
 Db 2 PIGNLGNPNVNSIPAPPLPSQTDGA 29

RESULT 6
 AAY06221
 ID AAY06221 standard; protein; 559 AA.
 AC AAY06221;
 XX
 DT 16-AUG-1999 (first entry)
 XX
 DE EHEC E. coli translocated intimin receptor (Tir).
 XX Tir; translocated intimin receptor; Hp90; enterohaemorrhagic; EHEC;
 KW infection; diagnosis; vaccine.
 XX
 OS Escherichia coli.
 XX
 FH Key Location/Qualifiers
 FT MISC-difference 453
 FT /note= "encoded by codon of 1 apparent nucleotide,
 FT causing frameshift in the DNA sequence"
 XX
 PN WO9924576-A1.
 XX
 PD 20-MAY-1999.
 XX
 PF 10-NOV-1998; 98WO-CA001042.
 XX
 PR 12-NOV-1997; 97US-0065130P.
 XX
 PA (UYBR-) UNIV BRITISH COLUMBIA.
 XX

PI Finlay BB, Kenny B, Devinney R, Stein M;
 XX
 DR WPI: 1999-337712/28.
 DR N-PSDB; AAX58859.
 XX
 PT New translocated intimin receptor useful for treating infection by
 PT enteropathogenic or enterohaemorrhagic Escherichia coli.
 XX
 PS Claim 7; Page 55-58; 91pp; English.
 XX
 CC The present sequence represents Tir, a novel translocated intimin
 CC receptor (formerly termed Hp90) from an enterohaemorrhagic Escherichia
 CC coli (EHEC) strain. The sequence was deduced from an isolated tir
 CC polynucleotide (see AAX58859). Tir proteins are secreted by attaching and
 CC effacing pathogens such as EHEC and EPEC (see AAY06220) E. coli. The
 CC bacterial pathogens insert their own receptors into mammalian cell
 CC surfaces, to which the pathogen then adheres to trigger additional host
 CC signaling events and actin nucleation. Diagnosis of disease caused by
 CC pathogenic E. coli can be performed by use of antibodies that bind to Tir
 CC to detect the protein or the use of nucleic acid probes for detection of
 CC nucleic acids encoding Tir polypeptide. Isolated Tir nucleic acids, Tir
 CC peptides, a recombinant method for producing recombinant Tir, antibodies
 CC which bind to Tir, and a kit for the detection of Tir-producing E. coli
 CC are provided. A method of immunising a host with Tir to induce a
 CC protective immune response is also provided. In addition, Tir fusion
 CC proteins can be used in attenuated E. coli to induce a cell-mediated
 CC immune response to other polypeptides, e.g. antigens. A method for
 CC screening for compounds which interfere with the binding of bacterial
 CC pathogens to their receptors is further provided
 CC
 SQ Sequence 559 AA;

Query Match 74.5%; Score 123; DB 2; Length 559;
 Best Local Similarity 82.1%; Pred. No. 7.3e-07;
 Matches 23; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 QY 1 PIGNLGNVNGNHLIPAPPLPSQTDGA 28
 |||||:|:|||||||
 Db 2 PIGNLGNPNVNSIPAPPLPSQTDGA 29

RESULT 7
 ABB60827
 ID ABB60827 standard; protein; 613 AA.
 AC ABB60827;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 9273.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 XX
 OS Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US009231.
 XX
 PR 23-MAR-2000; 2000US-0191637P.
 XX
 PR 11-JUL-2000; 2000US-00614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 XX
 DR WPI: 2001-656860/75.
 DR N-PSDB; ABL04930.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more

PT genes from *Drosophila* and for elucidating cell signaling and cell-cell
PT interactions.
XX
XX
PS Disclosure; SEQ ID NO 9273; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from *Drosophila*. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
CC sequences (AB101840-AB16175) and the encoded proteins (AB57737-
CC AB82072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
CC
SQ Sequence 613 AA;

Query Match 36.4%; Score 60; DB 4; Length 613;
Best Local Similarity 47.8%; Pred. No. 64;
Matches 11; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 2 IGNLNNVNGNHLIPAPLPSPQ 24
DB 449 IGMTGNQLNSLSDLPSPVPPDQ 471

RESULT 8
AEA26275 standard; protein; 329 AA.
AC AEA26275;
XX
XX 28-JUL-2005 (first entry)
DE Stress tolerant plant-related transcription factor protein SeqID116.
XX
XX transcription factor; transgenic plant; agriculture; drought resistance;
XX stress tolerance.
XX
XX *Oryza sativa*.
XX
XX WO2005047516-A2.
XX
XX 26-MAY-2005.
XX
XX 12-NOV-2004; 2004WO-US037584.
XX
XX 13-NOV-2003; 2003US-00714887.
XX
XX 05-DEC-2003; 2003US-0527658P.
XX
XX 05-FEB-2004; 2004US-0542928P.
XX
XX (MEND-) MENDEL BIOTECHNOLOGY INC.
XX
XX
XX Heard JR, Riechmann JL, Creelman RA, Ratcliffe OJ, Canales RD,
XX Repetti P, Kumimoto RW, Guttersen NI, Reuber TL, Pineda O, Adam L,
XX Sherman BK, Morrison TA, Keddie JS, Jiang C, Century KS, Adam L,
XX Zhang JZ, Hempel PD, Libby JW;
XX
XX WPI; 2005-372386/38.
XX
XX N-PSDB; AEA26274.
XX
XX
XX New transgenic plants for producing commercially or agriculturally useful
XX plants having improved tolerance to drought, shade and low nitrogen
XX conditions.
XX
XX Example 8; SEQ ID NO 116; 407pp; English.
XX
XX This invention relates to a novel plant transcription factor
XX polypeptides, the DNA sequences which encode them and their use in
XX creating transgenic plants. The transgenic plant and methods are useful
XX for producing commercially or agriculturally useful plants having
XX improved tolerance to drought, shade and low nitrogen conditions when

CC compared to wild-type reference plants. The present sequence is that of a
CC plant transcription factor protein which was used during the development
CC of the transgenic plants of the invention.
CC
SQ Sequence 329 AA;

Query Match 35.8%; Score 59; DB 9; Length 329;
Best Local Similarity 38.5%; Pred. No. 45;
Matches 10; Conservative 4; Mismatches 12; Indels 0; Gaps 0;

QY 3 GNLGNVNGNHLIPAPLPSPQTDGA 28
DB 35 GGGGGGAGNRHSPRPRAAESKA 60

RESULT 9
ABP41697 standard; protein; 604 AA.
ID ABP41697
XX
XX ABP41697;
XX
XX 22-AUG-2002 (first entry)
XX
XX
XX Human ovarian antigen H88MH38, SEQ ID NO:2829.
XX
XX
XX Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
XX ovarian cancer; breast cancer; tumour; reproductive system disorder;
XX infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
XX PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;
XX inflammatory condition; immune disorder; blood disorder;
XX cardiovascular disorder; respiratory disorder; neurological disorder;
XX gastrointestinal disorder; urinary system disorder; drug screening;
XX gene therapy; chromosome mapping; forensic analysis;
XX antibody preparation; cytostatic; immunomodulatory; neuroprotective;
XX antiinflammatory; gynaecological; reproductive.
XX
XX Homo sapiens.
XX
XX
XX WO200200677-A1.
XX
XX 03-JAN-2002.
XX
XX 07-JUN-2001; 2001WO-US018569.
XX
XX 07-JUN-2000; 2000US-0209467P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX
XX Birse CE, Rosen CA;
XX
XX WPI; 2002-147878/19.
XX
XX N-PSDB; ABQ54774.
XX
XX
XX Isolated nucleic acid molecules encoding novel ovarian polypeptides,
XX useful in the prevention, treatment and diagnosis of cancer (e.g. ovarian
XX cancer), immune disorders, cardiovascular disorders and neurological
XX diseases.
XX
XX Claim 11; SEQ ID NO 2829; 2922pp; English.
XX
XX
XX The invention relates to 2175 novel human ovarian antigens (ABP41054-
XX ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
XX encompasses polypeptides 90% identical and polynucleotides 95% identical
XX to the sequences of the invention. The invention additionally relates to
XX recombinant vectors and host cells comprising human ovarian antigen
XX polynucleotides, antibodies against human ovarian antigens, and the use
XX of ovarian antigen polynucleotides and polypeptides in diagnosing,
XX treating, prognosing or preventing various ovary and/or breast-related
XX disorders. Such conditions include ovarian cancer and breast cancer, and
XX metastatic tumours of ovarian or breast origin, reproductive system
XX disorders (e.g., infertility, disorders of pregnancy, anovulation,
XX polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine
XX disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic

PF 29-JUN-2001; 2001WO-US021067.
 XX
 PR 30-JUN-2000; 2000US-0215454P.
 PR 18-JUL-2000; 2000US-0219462P.
 PR 12-OCT-2000; 2000US-0240106P.
 PR 12-OCT-2000; 2000US-0240111P.
 PR 27-OCT-2000; 2000US-0244021P.
 PR 14-NOV-2000; 2000US-0248887P.
 PR 16-NOV-2000; 2000US-0249570P.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Tang YT, Yue H, Azimzai Y, He A, Batra S, Lo TP, Nguyen DB;
 PI Burdill JD, Marcus GA, Zingler KA, Gandhi AR, Lal P, Kearney L;
 PI Burford N, Yao MG, Walla NK, Elliot VS, Patterson C, Khan FA,
 PI Baughn MR, Hatalia AJA, Policky JL, Au-Young J, Lu Y, Borowsky ML,
 PI Lu DM, Ramkumar J, Yang J, Gururajan R, Warren BA, Gietzen K,
 PI Xu Y, Kallick DA, Lee EA, Thangavelu K, Delegeane AM, Lee S;
 XX
 DR WPI: 2002-154732/20.
 DR N-PSDB; ADI28057.
 XX
 PT Novel isolated human extracellular matrix and cell adhesion molecules
 PT useful for treating, preventing connective tissue disorder e.g. cystic fibrosis,
 PT osteoporosis, Paget's disease, genetic disorder e.g. cystic fibrosis,
 PT thalassemia.
 XX
 PS Claim 1; SEQ ID NO 13; 270pp; English.
 XX
 CC The invention relates to a novel isolated human extracellular matrix and
 CC cell adhesion molecule (referred to as ECMCAD 1-36), its biologically
 CC active or immunogenic fragment or a sequence comprising 90 % identity to
 CC ECMCAD 1-36. The molecule is useful for screening a compound for
 CC effectiveness as agonist or antagonist of itself. The protein and its
 CC encoding nucleic acid are useful in the diagnosis, treatment and
 CC prevention of genetic disorder such as adrenoleukodystrophy, Down's
 CC syndrome, cystic fibrosis, Gaucher's disease, myotonic dystrophy, sickle
 CC cell anemia, thalassemia, Wilms' tumor, etc, immunological disorders such
 CC as acquired immunodeficiency syndrome (AIDS), adult respiratory distress
 CC syndrome, allergies, anemia, asthma, atherosclerosis, autoimmune
 CC hemolytic anemia, contact dermatitis, Goodpasture's syndrome, gout,
 CC Grave's disease, multiple sclerosis, osteoporosis, psoriasis, rheumatoid
 CC arthritis, scleroderma, systemic lupus erythematosus, ulcerative colitis,
 CC uveitis, etc, a neurological disorder such as epilepsy, stroke, multiple
 CC Alzheimer's disease, Huntington's disease, Parkinson's disease, multiple
 CC sclerosis, bacterial and viral meningitis, periodic paralysis, mental
 CC disorders including mood, anxiety, and schizophrenic disorders, amnesia,
 CC diabetic neuropathy, etc, connective tissue disorder such as
 CC osteoporosis, Paget's disease, osteonecrosis, osteomyelitis,
 CC chondrosarcoma, giant cell tumor, psoriatic arthritis, infectious
 CC arthritis, systemic sclerosis, etc, and a cell proliferative disorder
 CC such as actinic keratosis, atherosclerosis, hepatitis, psoriasis, cancers
 CC including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, etc.
 CC This sequence represents one of the novel proteins.
 XX
 SQ Sequence 987 AA;
 XX
 QY Query Match 35.5%; Score 58.5; DB 5; Length 987;
 QY Best Local Similarity 44.4%; Pred. No. 1.6e+02;
 QY Matches 12; Conservative 5; Mismatches 7; Indels 3; Gaps 1;
 DB 2 IGNLGNVNGNHLIPAPLPQSOTDGA 28
 912 VGGMPFSVYGNMIPVAPIP---DGA 935
 RESULT 12
 ID AAM41013 standard; protein; 1086 AA.
 XX AAM41013;
 AC AAM41013;
 XX
 DT 22-OCT-2001 (first entry)

XX
 DE Human polypeptide SEQ ID NO 5944.
 XX
 KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemoatrophic;
 KW chemoatrophic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia.
 XX
 OS Homo sapiens.
 OS
 PN WO200153312-A1.
 XX
 PD 26-JUL-2001.
 XX
 PF 26-DEC-2000; 2000WO-US034263.
 XX
 PR 23-DEC-1999; 99US-00471275.
 PR 21-JAN-2000; 2000US-00488725.
 PR 25-APR-2000; 2000US-00552317.
 PR 20-JUN-2000; 2000US-00598042.
 PR 19-JUL-2000; 2000US-00620312.
 PR 03-AUG-2000; 2000US-00653450.
 PR 14-SEP-2000; 2000US-00662191.
 PR 19-OCT-2000; 2000US-00693036.
 PR 29-NOV-2000; 2000US-00727344.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang X, Weinman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
 PI Zhou P, Goodrich R, Dimaac RT;
 XX
 DR WPI: 2001-442253/47.
 DR N-PSDB; AAI60169.
 XX
 PT Novel nucleic acids and polypeptides, useful for treating disorders such
 PT as central nervous system injuries.
 XX
 PS Example 2; SEQ ID NO 5944; 10078pp; English.
 XX
 CC The invention relates to human nucleic acids (AA157798-AA161369) and the
 CC encoded polypeptides (AAM38642-AA42213) with nootropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localized neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemia and
 CC C.N.S disorders. Note: The sequence data for this patent did not form
 CC part of the printed specification
 XX
 SQ Sequence 1086 AA;
 XX
 QY Query Match 35.5%; Score 58.5; DB 4; Length 1086;
 QY Best Local Similarity 44.4%; Pred. No. 1.8e+02;
 QY Matches 12; Conservative 5; Mismatches 7; Indels 3; Gaps 1;
 DB 2 IGNLGNVNGNHLIPAPLPQSOTDGA 28
 1011 VGGMPFSVYGNMIPVAPIP---DGA 1034
 RESULT 13
 ID ADQ89830 standard; protein; 2486 AA.
 XX ADQ89830
 XX

AC ADQ89830;
XX
DT 21-OCT-2004 (first entry)
XX
DE Antagonist of cell cycle progression polypeptide #130.
XX
KW Cytostatic; cancer; cell division cycle; mitosis; meiosis;
KW cell cycle progression.
XX
OS Homo sapiens.
XX
PN WO2004063362-A2.
XX
PD 29-JUL-2004.
XX
PE 31-DEC-2003; 2003WO-GB005635.
XX
PR 10-JAN-2003; 2003US-0439123P.
PR 06-MAY-2003; 2003US-0468402P.
XX
PA (CYCL-) CYCLACEL LTD.
XX
PI Glover D, Bell G, Frenz L, Midgley C,
XX
XX WPI; 2004-544089/52.
DR N-PSDB; ADQ89829.
XX
PT New cell cycle progression genes and proteins for modulating cell cycle
PT progression in cells, for preventing, treating or diagnosing cell
PT proliferative diseases (e.g. cancer) or for identifying modulators of
PT mitosis or meiosis.
XX
PS Claim 2; SEQ ID NO 260; 461bp; English.
XX
CC The present invention relates to a polynucleotide for preventing,
CC treating or diagnosing a disease in an individual. The composition or the
CC polypeptide, polynucleotide or RNA precursor, or antibody is useful for
CC diagnosing, preventing or treating diseases (e.g. cell proliferative
CC diseases such as cancer) in an individual. These may also be used for
CC identifying substances capable of binding to or modulating the function
CC of the polypeptide, capable of affecting the function of the
CC corresponding gene, or capable of inhibiting the cell division cycle or
CC cell cycle progression, preferably mitosis and/or meiosis. The present
CC sequence represents an antagonist of cell cycle progression protein
CC sequence.
XX
SQ Sequence 2486 AA;
XX
Query Match 35.5%; Score 58.5; DB 8; Length 2486;
Best Local Similarity 44.4%; Pred. No. 4.3e+02;
Matches 12; Conservative 5; Mismatches 7; Indels 3; Gaps 1;
OY 2 I G N I G N N V N G N H L I P P A P L P S Q T D G A 28
DB 2411 V G G M P F S Y V G N A M I P P V A P I P --- D G A 2434
RESULT 14
ADX07462
ID ADX07462 standard; protein; 2602 AA.
XX
AC ADX07462;
XX
DT 21-APR-2005 (first entry)
XX
DE Cyclin-dependent kinase modulation biomarker SEQ ID NO 2027.
XX
KW Cytostatic; cyclin-dependent kinase; cdk; biomarker.
XX
OS Homo sapiens.
XX
PN WO2005012875-A2.
XX

PD 10-FEB-2005.
XX
XX 29-JUL-2004; 2004WO-US024424.
XX
PR 29-JUL-2003; 2003US-0490890P.
XX
PA (BRIM) BRISTOL-MYERS SQUIBB CO.
XX
PI Li M, Rupnow BA, Webster KR, Jackson DG, Wong TW;
XX
XX WPI; 2005-163068/17.
DR N-PSDB; ADX07461.
XX
PT Biomarkers useful for predicting or determining the response of a mammal
PT to a cancer treatment comprising administration of a modulator of cyclin-
PT dependent kinase activity.
XX
PS Claim 5; SEQ ID NO 2027; 141bp; English.
XX
XX
CC This invention describes a novel method of predicting or determining
CC whether a mammal will respond or is responding to an anti-cancer agent
CC that modulates cyclin-dependent kinase (cdk) activity. The method
CC comprises measuring the level of one or more biomarkers selected from
CC 2774 biomarkers given in the specification (nucleotide sequence SEQ ID
CC NO:1246 (Genbank EST W28729) is especially preferred). The method of the
CC invention is utilized in a kit for determining or predicting whether
CC patient would be susceptible or resistant to treatment by an agent
CC modulating cdk activity. The invention also describes a method for
CC utilizing individualized genetic profiles for treating diseases and
CC disorders based on patient's response and molecular level, specialized
CC microarrays comprising the biomarkers described, antibodies directed
CC against the biomarkers and a cell culture model to identify biomarkers.
CC The cdk modulator is preferably N-5-[(5-(1,1-dimethylethyl)-2-
CC oxazolyl)methyl]thio]-2-thiazolyl-4-piperidine carboxamide, 0.5-L-
CC tartaric acid salt. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pat_sequences. This
CC sequence represents a biomarker used in the method of the invention.
XX
SQ Sequence 2602 AA;
XX
Query Match 35.5%; Score 58.5; DB 9; Length 2602;
Best Local Similarity 44.4%; Pred. No. 4.5e+02;
Matches 12; Conservative 5; Mismatches 7; Indels 3; Gaps 1;
OY 2 I G N I G N N V N G N H L I P P A P L P S Q T D G A 28
DB 2527 V G G M P F S Y V G N A M I P P V A P I P --- D G A 2550
RESULT 15
ABP73854
ID ABP73854 standard; protein; 809 AA.
XX
AC ABP73854;
XX
DT 30-JAN-2003 (first entry)
XX
DE Candida albicans essential protein SEQ ID NO 7691.
XX
KW Fungus; yeast; tetracyclin; promoter; GRACE strain; biosynthesis;
KW signal transduction; DNA replication; cell division; growth;
KW proliferation; Candida albicans; fungicide; antifungal.
XX
OS Candida albicans.
XX
PN WO200253728-A2.
XX
PD 11-JUL-2002.
XX
PF 26-DEC-2001; 2001WO-US049486.
XX
PR 29-DEC-2000; 2000US-0259128P.
XX

PS Disclosure; Fig 12A; 0pp; English.

CC The invention relates to a method of diagnosing recurrent pregnancy loss
 CC (RPL). The method involves examining formin (Fmn)-2 gene for a mutation
 CC and measuring biological activity and expression of Fmn-2, in which
 CC decreased levels indicates an increased risk for RPL; or examining the
 CC person's formin-2 gene for polymorphisms, in which the presence of a
 CC polymorphism indicates an altered risk for RPL. The method is used for
 CC diagnosing and treating RPL e.g. in humans. The present sequence is mouse
 CC Fmn-2 protein
 XX

SQ Sequence 1567 AA;

Query Match 35.2%; Score 58; DB 7; Length 1567;

Best Local Similarity 57.9%; Pred. No. 3e+02; Mismatches 2; Indels 4; Gaps 1;

Matches 11; Conservative 2; Mismatches 2; Indels 4; Gaps 1;

OY 4 NLGNNGNHLIPAPPLP 22

DB 831 SFGNNCN---VPPAPPLP 845

RESULT 18

ABBS8090 ID ABB8090 standard; protein; 241 AA.

AC ABB8090;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 1062.

KW Drosophila; developmental biology; cell signalling; insecticide;

KW pharmaceutical.

OS Drosophila melanogaster.

PN WO200171042-A2.

PD 27-SEP-2001.

PE 23-MAR-2001; 2001WO-US009231.

PR 23-MAR-2000; 2000US-0191637P.

PR 11-JUL-2000; 2000US-00614150.

PA (PEKE) PE CORP NY.

PI Venter JC, Adams M, Li PWD, Myers EW;

DR WPI; 2001-656860/75.

DR N-PSDB; ABL02193.

PT New isolated nucleic acid detection reagent for detecting 1000 or more

PT genes from Drosophila and for elucidating cell signaling and cell-cell

PT interactions.

PS Disclosure; SEQ ID NO 1062; 21pp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent

CC capable of detecting 1000 or more genes from Drosophila. The invention is

CC useful in developmental biology and in elucidating cell signaling and

CC cell-cell interactions in higher eukaryotes for the development of

CC insecticides, therapeutics and pharmaceutical drugs. The invention

CC discloses genomic DNA sequences (ABL16176-ABU30511), expressed DNA

CC sequences (ABU01840-ABU16175) and the encoded proteins (ABB57737-

CC ABB72072). The sequence data for this patent did not form part of the

CC printed specification, but was obtained in electronic format directly

CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 241 AA;

Query Match 34.5%; Score 57; DB 4; Length 241;

Best Local Similarity 48.1%; Pred. No. 58;

Matches 13; Conservative 2; Mismatches 10; Indels 2; Gaps 1;

OY 3 NLGNNGNHL--IPAPPLPSQTDG 27

DB 84 GKNGFQASGDHLPOAPAPLPQPVPTAG 110

RESULT 19

AAW05424 ID AAW05424 standard; peptide; 22 AA.

AC AAW05424;

DT 24-FEB-1998 (first entry)

DE SH3 domain peptide recognition unit pCkr.

KW Src-homology region 3 domain; human; mouse; SH3 domain; cell growth;

KW cellular signalling element; cellular structural element; malignancy;

KW protein identification; functional domain; protein screening;

KW cellular signal transduction process; binding peptide.

OS Homo sapiens.

PN WO9631625-A1.

PD 10-OCT-1996.

PE 04-APR-1996; 96WO-US004454.

PR 07-APR-1995; 95US-00417872.

PR 03-APR-1996; 96US-00630915.

PA (CYTO-) CYTOGEN CORP.

PA (UNNC-) UNIV NORTH CAROLINA.

PI Sparks AB, Hoffman N, Kay BK, Fowles DM, McConnell SJ;

DR WPI, 1996-465045/46.

PT Identifying polypeptide(s) having specific functional domain (esp. SH3

PT domain) - comprises detecting selective binding to recognition unit,

PT regardless of sequence homology.

PS Example; Page 87; 174pp; English.

CC AAW05421-W05424 represent Src-homology region 3 (SH3) domain peptide

CC recognition units. These sequences were used as parts of multivalent

CC recognition unit complexes used in the method of the invention. The

CC method of the invention is for identifying polypeptides containing

CC functional domains of interest (especially SH3 domains). It comprises

CC contacting a multivalent recognition unit (RU) complex with a number of

CC peptides and identifying polypeptides having a selective binding affinity

CC for the RU complex. The method is based on functional similarities and

CC does not rely on sequence similarities. Prior methods only gave limited

CC success for identifying proteins containing an SH3 domain due to the

CC minimal sequence homology among known SH3 proteins. Multivalent RU

CC containing functional domains that are similar to, but not identical in

CC sequence to, the original target functional domain. The new method

CC enables proteins having a common function to be identified.

CC Identification of novel SH3 proteins will be useful for a better

CC understanding of cell growth, malignancy, signal transduction processes,

CC etc. New candidate drugs can be identified, and their specificities (e.g.

CC pharmacological activities) can be assessed using the method of the

XX invention

XX Sequence 22 AA;

Query Match 34.2%; Score 56.5; DB 2; Length 22;
 Best Local Similarity 50.0%; Pred. No. 5.5;
 Matches 11; Conservative 3; Mismatches 5; Indels 3; Gaps 1;

QY 6 GNNVNGNHLPPAPPLPSQTDG 27
 |||||:|||||:
 DB 4 GNYVNA---LPPGPPLPAKNGG 22

RESULT 20

ABP08155
 ID ABP08155 strand; protein; 134 AA.

XX AC ABP08155;

XX DT 24-JUN-2002 (first entry)

XX DE Human ORFX protein sequence SEQ ID NO.16292.

XX KW Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;
 KW hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;
 KW degenerative disorder; osteoarthritis; neurodegenerative disorder;
 KW cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
 KW hypertension; hypothyroidism; cholesterol ester storage disease;
 KW immune deficiency; immune disorder; infectious disease;
 KW autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
 KW myasthenia gravis.

XX OS Homo sapiens.

XX PN WO200192523-A2.

XX PD 06-DEC-2001.

XX PF 29-MAY-2001; 2001WO-US010836.

XX PR 30-MAY-2000; 2000US-0206132P.

XX PR 29-AUG-2000; 2000US-0228716P.

XX PA (CURA-) CURAGEN CORP.

XX PI Shinketsu RA, Leach MD;

XX DR WPI: 2002-106308/14.

XX DR N-PSDB; ABN233907.

XX PT Novel human polypeptides and polynucleotides useful for diagnosing,
 PT preventing and treating cardiovascular disease, neurodegenerative,
 PT hyperproliferative disorders and autoimmune disorders.

XX PS Disclosure; SEQ ID NO 16292; 1037pp; English.

XX CC The present invention describes substantially purified human proteins
 CC (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1
 CC in the specification). ABN15762 to ABN27252 encode the human ORFX
 CC proteins given in ABP00010 to ABP1500. ORFX proteins are useful for
 CC treating or preventing a pathology associated with an ORFX-associated
 CC syndrome in humans, and in the manufacture of a medicament for treating a
 CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide
 CC sequences can be used in gene therapy. ORFX sequences can be used in the
 CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,
 CC psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,
 CC osteoarthritis, neurodegenerative disorders, disorders related to organ
 CC transplantation, cardiovascular diseases, diabetes mellitus, systemic
 CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester
 CC storage disease, various immune deficiencies and disorders, infectious
 CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid
 CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host
 CC disease and autoimmune inflammatory eye disease. ORFX proteins are also
 CC useful for treating burns, incisions, ulcers, for treating osteoporosis,
 CC bone degenerative disorders, or periodontal disease, and for gut
 CC protection or regeneration and treatment of lung or liver fibrosis.

CC reperfusion injury in various tissues and conditions resulting from
 CC systemic cytokine damage. N.B. The sequence data for this patent did not
 CC form part of the printed specification, but was obtained in electronic
 CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX SQ Sequence 134 AA;

QY Query Match 34.2%; Score 56.5; DB 5; Length 134;
 Best Local Similarity 46.2%; Pred. No. 36;
 Matches 12; Conservative 1; Mismatches 12; Indels 1; Gaps 1;

QY 3 GNLGNNVNGNH-LPPAPPLPSQTDG 27
 |||||:|||||:
 DB 28 GRFGNGTNGDHPKGPDPPEDEKG 53

RESULT 21

ADJ95526
 ID ADJ95526 strand; peptide; 42 AA.

XX AC ADJ95526;

XX DT 06-MAY-2004 (first entry)

XX DE Equine herpesvirus type 1-related peptide #1.

XX KW virulence; herpesvirus; Equine herpesvirus type 1; EHV-1; EHV-4;
 KW Equine herpesvirus type 4; genetic marker; ORF30-ml region marker;
 KW vaccine; EHV-1 strain V592 polymerase ORF30-ml region;
 KW neurovirulence; herpesvirus disease.

XX OS Equine herpesvirus 1.

XX PN WO2004011677-A2.

XX PD 05-FEB-2004.

XX PF 23-JUL-2003; 2003WO-GB003279.

XX PR 26-JUL-2002; 2002US-0398576P.

XX PA (ANIM-) ANIMAL HEALTH TRUST.

XX PI Davis Poynter N, Nugent J, Birch-Machin I, Allen G;

XX DR WPI: 2004-143877/14.

XX PT Assessing the virulence of a herpesvirus isolate, useful in preventing or
 PT treating herpesvirus infection, by using virulence marker corresponding
 PT to an ORF30-ml region.

XX PS Disclosure; Page 43; 63pp; English.

XX CC This invention relates to a novel method of assessing the virulence of a
 CC herpesvirus isolate, for example Equine herpesvirus type 1 (EHV-1) or
 CC type 4 (EHV-4), which comprises using a genetic marker, especially an
 CC ORF30-ml region marker. The invention may be useful for the production of
 CC compounds with a virulence activity or for the development of a vaccine.
 CC The method, kit and sequence of the EHV-1 strain V592 polymerase ORF30-ml
 CC region are useful for assessing the virulence of a herpesvirus or
 CC neurovirulence. The vaccine is useful for immunising a host against a
 CC herpesvirus disease and for treating disease. The present sequence is
 CC that of a peptide which is related to the invention.

XX SQ Sequence 42 AA;

QY Query Match 33.9%; Score 56; DB 8; Length 42;
 Best Local Similarity 66.7%; Pred. No. 13;
 Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 15 IPPAPPLPSQTDGAA 29
 :|||||:|||||:
 DB 3 LPPAPPLPOSTSKAA 17

```

RESULT 22
AAB32745 ID AAB32745 standard; protein; 145 AA.
XX
XX
AC AAB32745;
XX
XX DT 25-JAN-2001 (first entry)
XX
DE Eucalyptus grandis transcription factor protein sequence #203.
XX
XX Plant; transcription factor; gene expression; eucalyptus; pine; acacia;
KW poplar; sweetgum; teak; mahogany; bZIP; G-box binding factor;
KW basic helix-loop-helix zipper; homeotic; homeodomain; homeobox; MADS;
KW homeodomain zipper; LIM domain; AP2; ERBBs; zinc finger domain;
KW type 2 Cys2His2; CCAAT box element; MYB.
XX
XX Eucalyptus grandis.
XX
XX OS MO200053724-A2.
XX
XX PN 14-SEP-2000.
XX
XX PF 09-MAR-2000; 2000WO-US006112.
XX
XX PR 11-MAR-1999; 99US-00266513.
XX
XX PR 18-AUG-1999; 99US-0149485P.
XX
XX (GENE-) GENESIS RES & DEV CORP LTD.
XX
XX (FLET-) FLETCHER CHALLENGE FORESTS LTD.
XX
XX PI Wood M, McGrath A, Shenk MA, Glenn M;
XX
XX WPI: 2000-579369/54.
XX
XX DR New isolated polynucleotide encoding a plant transcription factor for
XX
XX PT producing a plant e.g. a woody plant, preferably eucalyptus or pine,
XX
XX PT having modified gene expression or modified activity of a polypeptide.
XX
XX Claim 8; Page 284-285; 747pp; English.
XX
XX The present invention relates to novel plant transcription factors from
XX
XX CC Eucalyptus grandis or Pinus radiata. The present sequence is one such
XX
XX CC transcription factor. The transcription factor may be used to produce a
XX
XX CC plant having modified gene expression such as a woody plant e.g. a
XX
XX CC eucalyptus, pine, acacia, poplar, sweetgum, teak, or mahogany species or
XX
XX CC to modify the activity of a polypeptide in a plant. The transcription
XX
XX CC factors of the present invention are members from the following families
XX
XX CC of regulatory proteins: bZIP, bZIP family of G-box binding factors, basic
XX
XX CC helix-loop-helix zipper, homeotic/homeodomain/homeobox/MADS, homeodomain
XX
XX CC zipper, LIM domain, AP2 and ERBBs, zinc finger domains of type 2
XX
XX CC Cys2His2, CCAAT box elements and MYB
XX
XX SQ Sequence 145 AA;

Query Match 33.9%; Score 56; DB 3; Length 145;
Best Local Similarity 38.5%; Pred. No. 45;
Matches 10; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

```

```

XX
XX Shrump white spot Bacilliform virus (WSBV) protein 42.
DE
XX
XX Shrimp white spot Bacilliform virus; WSBV; diagnosis; viral infection;
KW antiviral agent; gene expression; antisense construct;
KW transgenic viral resistant shrimp.
XX
XX OS White spot syndrome virus.
XX
XX PN WO200138351-A2.
XX
XX PD 31-MAY-2001.
XX
XX PF 08-NOV-2000; 2000WO-US028888.
XX
XX PR 24-NOV-1999; 99CN-00124717.
XX
XX (PENY-) PE CORP NY.
XX
XX PA (THIR-) THIRD INST OCEANOGRAPHY STATE OCEANI C A.
XX
XX PA (SINO-) SINOGENOMAX CO LTD.
XX
XX PI Xu X, Yang F, He J, Pham L, He M, Ye Y, Shen Y, Kodira C;
XX
XX DR WPI: 2001-355877/37.
XX
XX DR N-PSDB; AAH62731.
XX
XX PT Primary nucleotide sequence of the shrimp white spot Bacilliform virus
XX
XX PT (WSBV), useful for producing viral polypeptides that can be used to
XX
XX PT screen for agents that are useful for treating WSBV infection.
XX
XX Claim 1; Fig 3; 626pp; English.
XX
XX The invention provides the primary nucleotide sequence of the WSBV genome
XX
XX CC (AAH62689), predicted transcript sequences (AAH62689-AAH62839) and
XX
XX CC encoded proteins (AAH64910-AAH65051) and oligonucleotide sequences
XX
XX CC (AAH62840-63160) suitable for use as primers or probes. The nucleic acid
XX
XX CC molecules and proteins of the invention are useful for diagnosis and
XX
XX CC monitoring viral infection. In screens for antiviral agents and for
XX
XX CC monitoring viral gene expression or activity during a treatment regimen.
XX
XX CC The nucleic acid molecules are also useful as antisense constructs to
XX
XX CC control viral gene expression in infected cells and tissues and to create
XX
XX CC transgenic viral resistant shrimp. (Updated on 06-AUG-2003 to correct OS
XX
XX CC field.)
XX
XX SQ Sequence 295 AA;

Query Match 33.9%; Score 56; DB 4; Length 295;
Best Local Similarity 39.4%; Pred. No. 95;
Matches 13; Conservative 3; Mismatches 9; Indels 8; Gaps 1;

QY 1 PIGNLGNV-----VNGNHLIPAPPLPSQT 25
DB 124 PSVNTGNTGGGGGTGTVBGLPLPPPTPTPT 156

RESULT 24
AAB96146 ID AAB96146 standard; protein; 410 AA.
XX
XX
AC AAB96146;
XX
XX DT 29-OCT-2001 (first entry)
XX
XX DE Putative P. abyssi 5-enolpyruvylshikimate-3-phosphate synthase.
XX
XX KW Hyperthermophilic archaeon; hyperthermophilic protein.
XX
XX OS Pyrococcus abyssi.
XX
XX PN FR2792651-A1.
XX
XX PD 27-OCT-2000.
XX

```

PF 21-APR-1999; 99FR-00005034.
XX
PR 21-APR-1999; 99FR-00005034.
XX
PA (CNRS) CNRS CENT NAT RECH SCI.
PI (IFRE-) IFREMER INST FR RECH EXPL MER.
XX
PI Porterre P, Thiery JC, Prieur D, Dietrich J, Lecompte O;
PI Quereillon J, Weissenbach J, Saurin W, Heilig R;
XX WPI; 2001-126236/14.
DR
XX New nucleotide sequences isolated from *Pyrococcus abyssi* encode proteins
PT useful in industry.
XX
PS Claim 7; Page 783-784; 1657pp; French.
XX
CC The present invention relates to the genomic sequence of *Pyrococcus*
CC *abyssi* (see AA066431 and AA01223-7) and *P. abyssi* proteins. *P. abyssi* is
CC a hyperthermophilic archaeon, which is isolated from deep-sea
CC hydrothermal vents. The present sequence is one such *P. abyssi* protein.
CC The proteins of the present invention have various potential industrial
CC uses, since the proteins are stable at very high temperatures, some up to
CC 110 degrees centigrade. Note: This patent is in the same patent family as
CC WO00065062, which contains additional sequences as shown in AAB99132-
CC AAB99143, AA075903-AA075920 and AA066436
XX
SQ Sequence 410 AA;
XX
Query Match 33.9%; Score 56; DB 4; Length 410;
Best Local Similarity 44.0%; Pred. No. 1.3e+02;
Matches 11; Conservative 2; Mismatches 12; Indels 0; Gaps 0;
QY 2 IGNLGNVNGNHLIPAPPLPSQTD 26
DB 62 ISKFGAQNNGNKIIPQELTPGKID 86
XX
RESULT 25
ID ADS43142 standard; protein; 410 AA.
XX
AC ADS43142;
XX
DT 02-DEC-2004 (first entry)
XX
DE Bacterial polypeptide #21572.
XX
KW Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmotic;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polypeptide.
XX
OS Bacteria.
XX
PN US2003233675-A1.
XX
PD 18-DEC-2003.
XX
PF 20-FEB-2003; 2003US-00369493.
XX
PR 21-FEB-2002; 2002US-0360039P.
XX
PA (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
XX
PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
PI

XX
DR WPI; 2004-061375/06.
XX
PT New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
PS Claim 1; SEQ ID NO 21572; 122pp; English.
XX
CC The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 410 AA;
XX
Query Match 33.9%; Score 56; DB 8; Length 410;
Best Local Similarity 44.0%; Pred. No. 1.3e+02;
Matches 11; Conservative 2; Mismatches 12; Indels 0; Gaps 0;
QY 2 IGNLGNVNGNHLIPAPPLPSQTD 26
DB 62 ISKFGAQNNGNKIIPQELTPGKID 86
XX
RESULT 26
ID ABO66765 standard; protein; 474 AA.
XX
AC ABO66765;
XX
DT 29-JUL-2004 (first entry)
XX
DE *Klebsiella pneumoniae* polypeptide seqid 13282.
XX
KW Recombinant expression vector; transcription regulatory element;
KW *Klebsiella pneumoniae* protein; antibacterial; Vaccine.
XX
OS *Klebsiella pneumoniae*.
XX
PN US6610836-B1.
XX
PD 26-AUG-2003.
XX
PF 27-JAN-2000; 2000US-00489039.
XX
PR 29-JAN-1999; 99US-0117747P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Breton GL, Osborne M;
XX
DR WPI; 2003-895346/82.
XX

CC resistance to plant disease, for galactomannan production, for production
CC of plant growth regulators, for improving plant heat tolerance, for
CC improving plant tolerance to herbicides, for increasing the rate of
CC homologous recombination in plants, for lignin production, for improving
CC plant tolerance to extreme osmotic conditions, for improving plant
CC tolerance to pathogens or pests, for yield improvement by modification of
CC photosynthesis, for modifying seed oil yield and/or content, for
CC modifying seed protein yield and/or content, for yield improvement by
CC modification of carbohydrate, nitrogen or phosphorus use and/or uptake
CC and for yield improvement by providing improved plant growth and
CC development under at least one stress condition. The polynucleotide may
CC also encode a plant transfection factor. The methods and compositions of
CC the present invention are useful in the field of biochemistry and
CC genetics, in particular for producing transgenic plants with improved
CC biological characteristics such as increased yield, improved nitrogen
CC flow, increasing plant tolerance to cold or heat, improving plant
CC tolerance to extreme osmotic and drought conditions, and improving plant
CC tolerance to plant pests or pathogens. They can also be used in physical
CC arrays of molecules, plant breeding markers, computer-based storage and
CC analysis systems. The present sequence is one of the 5544 plant protein
CC sequences of the invention. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?docID=20040216190.
XX
SQ Sequence 301 AA;

Query Match 33.6%; Score 55.5; DB 8; Length 301;
Best Local Similarity 44.8%; Pred. No. 1.1e+02;
Matches 13; Conservative 2; Mismatches 9; Indels 5; Gaps 1;

QY 1 PIGNLGNVNGNHLT-----PPAPPLPSQ 24
Db 61 PPTPTTTNNNNNNLLIQNTNTPSPSPPPQ 89

RESULT 29
ID AEF21139 standard; protein; 627 AA.
XX AEF21139;
XX AEF21139;
DT 09-MAR-2006 (first entry)
XX
DE Candida albicans forin FRI-FR2 domains.
XX
KW High throughput screening; therapeutic; metastasis; cytosolic;
KW autoimmune disease; immunosuppressive; cardiovascular disease;
KW cardiovascular-gen.; blood clotting disorder; anticoagulant;
KW thrombolytic; hemostatic; forin.
XX
OS Candida albicans.
XX
PN US2006003399-A1.
XX
PD 05-JAN-2006.
XX
PF 10-JUN-2005; 2005US-00150845.
XX
PR 10-JUN-2004; 2004US-0578949P.
PR 20-APR-2005; 2005US-0673444P.
XX
PA (CYTO-) CYTOKINETICS INC.
XX
XX Tomasevic N, Jia Z, Sakowicz R, Pierce D, Finer J;
XX WPI; 2006-077959/08.
XX
DR
XX
PT Screening a modulator of a component involved in actin polymerization,
PT nucleation promoting factor, and an upstream regulator.
XX
PS Example 11; SEQ ID NO 47; 106pp; English.

XX
XX The present invention provides a method for high throughput screening of
CC an agent for modulating the activity of a component involved in actin
CC polymerization. The method involves combining pyrene-globular actin
CC (pyrene-G-actin) or acrylodan-globular actin (acrylodan-G-actin), Arp2/3
CC complex, nucleation promoting factors (NPFs) and an upstream regulator in
CC the presence of an agent and detecting the fluorescence which is a
CC measure of the polymerization of pyrene-G-actin into pyrene-filamentous
CC actin (pyrene-F-actin) or acrylodan-G-actin into acrylodan-filamentous
CC actin (acrylodan-F-actin). NPFs include Wiskott-Aldrich Syndrome (WASP)
CC protein, a WASP homolog called N-WASP and a family of proteins called
CC suppressor of CAR (SCAR), also referred to as the WASP family verprolin
CC homologous (WAVE) proteins. The SCAR/WAVE family includes SCAR1/WAVE1
CC protein, SCAR2/WAVE2 protein and a SCAR3/WAVE3 proteins and the
CC regulators may be Cdc42 protein, Rac1 protein, RhoA protein, Rhoc
CC protein, Nck1 protein, Nck2 protein and phosphatidylinositol-1,4-
CC bisphosphate (PIP2). The invention is useful for treating metastatic
CC cancer, autoimmune diseases, cardiovascular and inflammatory diseases and
CC conditions associated with hyperactivity of platelets or increased risk
CC of blood clotting. The present sequence is the Candida albicans forin
CC FRI-FR2 domain used to construct E. coli plasmids which is used in
CC the purification of the Arp2/3 complex that can be used in polymerization
CC assays.
XX
SQ Sequence 627 AA;

Query Match 33.6%; Score 55.5; DB 10; Length 627;
Best Local Similarity 55.0%; Pred. No. 2.4e+02;
Matches 11; Conservative 4; Mismatches 4; Indels 1; Gaps 1;

QY 9 VNGS-VLIPAPPLPSQIDG 27
Db 120 LMGSGSVIPAPPLPPSSG 139

RESULT 30
ID ADT60207 standard; protein; 319 AA.
XX ADT60207;
XX ADT60207;
DT 13-JAN-2005 (first entry)
XX
DE Plant polypeptide, SEQ ID 10284.
XX
XX Plant; transgenic; cold tolerance; growth rate; drought tolerance;
KW disease resistance; galactomannan production; plant growth regulator;
KW heat tolerance; herbicide tolerance; lignin production;
KW extreme osmotic condition tolerance; pathogens resistance;
KW pest resistance; yield improvement; seed oil yield; seed protein yield.
XX
OS Viridiplantae.
XX
PN US2004216190-A1.
XX
PD 28-OCT-2004.
XX
PF 18-DEC-2003; 2003US-00739930.
XX
PR 28-APR-2003; 2003US-00424599.
PR 28-APR-2003; 2003US-00425115.
XX
PA (KOVA/) KOVALIC D K.
XX
XX Kovalic DK;
XX WPI; 2004-757369/74.
XX
DR
XX
PT New recombinant DNA constructs useful in the field of biochemistry and
PT genetics, and in particular for producing transgenic plants with improved
PT biological characteristics.
XX
PS Claim 2; SEQ ID NO 10284; 14pp; English.

XX US6551795-B1.
 PN
 XX
 PD 22-APR-2003.
 XX
 PF 18-FEB-1999; 99US-00252991.
 XX
 PR 18-FEB-1998; 98US-0074788P.
 XX
 PR 27-JUL-1998; 98US-0094190P.
 XX
 PA (GENO-) GENOME THERAPEUTICS CORP.
 XX
 PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;
 XX
 DR MPI; 2003-615309/58.
 DR N-PSDB; ABD11269.
 PT Novel isolated nucleic acid encoding *Pseudomonas aeruginosa* polypeptide,
 PT useful as molecular targets for diagnostics, prophylaxis and treatment of
 PT pathological conditions resulting from bacterial infection.
 XX
 PS Disclosure; SEQ ID NO 26444; 455bp; English.
 XX
 CC The invention relates to *Pseudomonas aeruginosa* polypeptides and the
 CC polynucleotides encoding them. The sequences are useful in diagnosis and
 CC therapy of pathological conditions, as molecular targets for diagnostics,
 CC prophylaxis and treatment of pathological conditions resulting from a
 CC bacterial infection, for evaluating a compound, such as a polypeptide,
 CC for the ability to bind a *P. aeruginosa* nucleic acid, as components of
 CC effective antibacterial targets, as targets for antibacterial drugs,
 CC including anti-*P. aeruginosa* drugs, as templates for recombinant
 CC production of *P. aeruginosa*-derived peptides or polypeptides, as target
 CC components for diagnosis and/or treatment of *P. aeruginosa*-caused
 CC infection, and in detection of *P. aeruginosa* sequences or other sequences
 CC of *Pseudomonas* species using biochip technology. Sequences AB067826-
 CC AB084396 represent *P. aeruginosa* polypeptides of the invention. Note: The
 CC sequence data for this patent did not form part of the printed
 CC specification but was obtained in electronic format from USPTO at
 CC seqdata.uspto.gov/sequence.html
 CC
 SQ Sequence 631 AA;
 Query Match 33.3%; Score 55; DB 7; Length 631;
 Best Local Similarity 55.6%; Pred. No. 2.8e+02;
 Matches 10; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 7 NNVNGNHLPPAPPLPSP 24
 Db 397 NSLPGNPTPTPTPLPVQ 414

RESULT 33
 ABM86049
 ID ABM86049 standard; protein; 899 AA.
 XX
 AC ABM86049;
 XX
 DT 02-JUN-2005 (first entry)
 XX
 DE Rice abiotic stress responsive polypeptide SEQ ID NO:4295.
 XX
 KM abiotic stress tolerance; transgenic plant; plant; cereal; agriculture.
 XX
 OS *Oryza sativa*.
 XX
 PN WO2003008540-A2.
 XX
 PD 30-JAN-2003.
 XX
 PF 21-JUN-2002; 2002WO-US019668.
 XX
 PR 22-JUN-2001; 2001US-0300112P.
 XX
 PR 24-AUG-2001; 2001US-0314662P.

PR 26-SEP-2001; 2001US-0325277P.
 PR 21-NOV-2001; 2001US-0332132P.
 XX
 XX (SYGN) SYNGENTA PARTICIPATIONS AG.
 PA
 XX
 PI Krepe J, Briggs SP, Cooper B, Glazebrook J, Goff SA, Katagiri F;
 PI Moughamer T, Provart N, Rieke D, Zhu T;
 XX
 DR MPI; 2003-248011/24.
 XX
 PT New stress-responsive nucleic acid, useful for altering the
 PT responsiveness of a plant, e.g. cereal, to an abiotic stress such as cold
 PT stress, salt stress or osmotic stress.
 XX
 PS Claim 1; SEQ ID NO 4295; 89pp; English.
 XX
 CC The invention relates to novel abiotic stress responsive polynucleotides
 CC and polypeptides. Also disclosed are vectors, expression cassettes, host
 CC cells, and plants containing such polynucleotides. Also disclosed are
 CC methods for using the polynucleotides and polypeptides to alter the
 CC responsiveness of a plant to abiotic stresses. The invention is useful in
 CC agriculture. The nucleic acid is useful for determining whether a test
 CC plant has been exposed to an abiotic stress condition. It is also useful
 CC for selecting an agent that alters abiotic stress regulated
 CC polynucleotide expression in a plant cell, and to identify a homolog or
 CC ortholog to an abiotic stress responsive polynucleotide. The nucleic acid
 CC molecule and the polypeptide encoded by it are useful in altering the
 CC responsiveness of a plant to an abiotic stress, such as cold stress, salt
 CC stress, osmotic stress or any of their combinations. The present sequence
 CC is used in the exemplification of the invention
 CC
 SQ Sequence 899 AA;
 Query Match 33.3%; Score 55; DB 7; Length 899;
 Best Local Similarity 63.6%; Pred. No. 4.1e+02;
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 12 NHIIPAPPLP 22
 Db 703 NHIIPQPPIP 713

RESULT 34
 ABB68251
 ID ABB68251 standard; protein; 283 AA.
 XX
 AC ABB68251;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE *Drosophila melanogaster* polypeptide SEQ ID NO 31545.
 XX
 KM *Drosophila*; developmental biology; cell signalling; insecticide;
 KM pharmaceutical.
 XX
 OS *Drosophila melanogaster*.
 XX
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US009231.
 XX
 PR 23-MAR-2000; 2000US-0191637P.
 XX
 PR 11-JUL-2000; 2000US-00614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 XX
 DR MPI; 2001-656860/75.
 DR N-PSDB; ABL12354.
 XX

PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from *Drosophila* and for elucidating cell signaling and cell-cell
PT interactions.
XX
XX
XX Disclosure; SEQ ID NO 31545; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from *Drosophila*. The invention is
CC useful in developmental biology and in elucidating cell signaling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
CC sequences (AB101840-AB16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 283 AA;

Query Match 33.0%; Score 54.5; DB 4; Length 283;
Best Local Similarity 45.8%; Pred. No. 1.4e+02;
Matches 11; Conservative 3; Mismatches 3; Indels 7; Gaps 1;

Qy 13 HLIPAPP-----LPSQTDGAA 29
||:||||| :|||
Db 166 HLLPAPPSPSYDQATTTPAETTCGA 189

RESULT 35
AB074627
ID AB074627 standard; protein; 618 AA.
XX
XX AC AB074627;
XX DT 29-JUL-2004 (first entry)
XX DE Pseudomonas aeruginosa polypeptide #6802.
XX
XX Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX OS Pseudomonas aeruginosa.
XX PN US6551795-B1.
XX PD 22-APR-2003.
XX PF 18-FEB-1999; 99US-00252991.
XX PR 18-FEB-1998; 98US-0074788P.
XX PR 27-JUL-1998; 98US-0094190P.
XX PA (GENO-) GENOME THERAPEUTICS CORP.
XX PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX WPI; 2003-615309/58.
XX DR N-PSDB; ABD08198.
XX
XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.
XX
XX Disclosure; SEQ ID NO 23373; 455pp; English.
XX
XX The invention relates to Pseudomonas aeruginosa polypeptides and the
CC polymucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a P. aeruginosa nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-P. aeruginosa drugs, as templates for recombinant
CC production of P. aeruginosa-derived peptides or polypeptides, as target

CC components for diagnosis and/or treatment of P. aeruginosa-caused
CC infection, and in detection of P. aeruginosa sequences or other sequences
CC of Pseudomonas species using biochip technology. Sequences AB067826-
CC AB084386 represent P. aeruginosa polypeptides of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html
XX
SQ Sequence 618 AA;

Query Match 33.0%; Score 54.5; DB 7; Length 618;
Best Local Similarity 48.0%; Pred. No. 3.2e+02;
Matches 12; Conservative 3; Mismatches 7; Indels 3; Gaps 1;

Qy 6 GNNVNGNHLIPAPPPLPSQTDGAAR 30
||:||||| :|||
Db 320 GNRLDGEHL--RPPHRTAGAQK 341

RESULT 36
ADC30973
ID ADC30973 standard; protein; 207 AA.
XX
XX AC ADC30973;
XX DT 18-DEC-2003 (first entry)
XX
XX DE Human novel polypeptide sequence, SEQ ID NO:1055.
XX
XX Human; diagnostic; drug screening; forensics; gene mapping;
XX biodiversity assessment; Parkinson's disease; Alzheimer's disease;
XX neurodegenerative diseases; anaemia; platelet disorder; wound; burns;
XX ulcers; osteoporosis; autoimmune disease; cancer;
XX molecular weight marker; food supplement; antiparkinsonian; nootropic;
XX neuroprotective; antianaemic; anticoagulant; thrombolytic; vulnerary;
XX antifuror; osteopathic; immunosuppressive; antiinflammatory; cyostatic;
XX gene therapy; chromosome 8.
XX OS Homo sapiens.
XX PN WO2003029271-A2.
XX PD 10-APR-2003.
XX PF 24-SEP-2002; 2002WO-US030474.
XX PR 24-SEP-2001; 2001US-0324631P.
XX PA (HYSE-) HYSEQ INC.
XX PI Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;
XX Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;
XX Haley-Vicente D, Drmanac RT;
XX WPI; 2003-371981/35.
XX DR N-PSDB; ADC30002.
XX
XX New polymucleotide and polypeptide useful for diagnosing, preventing or
PT treating conditions such as neurodegenerative diseases, anemia, platelet
PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
PT cancer.
XX
XX Claim 20; SEQ ID NO 1055; 1185pp; English.
XX
XX The invention relates to 971 novel human cDNA sequences (ADC29919-
CC ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The
CC invention also relates to nucleic acid sequences over 99% identical with
CC the novel human cDNAs. The invention additionally encompasses expression
CC vectors and host cells comprising a nucleic acid of the invention; the
CC recombinant production of a polypeptide of the invention; an antibody
CC against a polypeptide of the invention; a method of detecting
CC polymucleotides or polypeptides of the invention; and methods of
CC identifying a compound which binds to a polypeptide of the invention. The

CC invention further discloses methods of preventing, treating or
 CC ameliorating a medical condition; kits comprising polynucleotide probes
 CC and/or monoclonal antibodies for carrying out the methods of the
 CC invention; methods for the identification of compounds that modulate the
 CC expression or activity of the polynucleotide and/or polypeptide; and 767
 CC contig sequences corresponding to the cDNA sequences of the invention
 CC (ADCC1861-ADCC32627) and the polypeptides encoded by the contigs (ADCC32628
 CC -ADCC33394). The nucleic acids and polypeptides of the invention are
 CC useful in diagnostics, drug screening, forensics, gene mapping, in the
 CC identification of mutations responsible for genetic disorders or other
 CC traits, for assessing biodiversity, and in producing many other types of
 CC data and products dependent on DNA and amino acid sequences. They are
 CC also used for treating diseases such as Parkinson's disease, Alzheimer's
 CC disease and other neurodegenerative diseases, anaemia, platelet
 CC disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
 CC cancer. The nucleic acids may also be used as hybridisation probes or
 CC primers, and in the recombinant production of a protein. The polypeptides
 CC are also useful in generating antibodies, as molecular weight markers,
 CC and as food supplements. The present sequence represents a specifically
 CC claimed human polypeptide sequence of the invention. Note: The sequence
 CC data for this patent did not form part of the printed specification, but
 CC was obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

CC Sequence 207 AA;

Query Match 32.7%; Score 54; DB 7; Length 207;
 Best Local Similarity 37.8%; Pred. No. 1.2e+02;
 Matches 14; Conservative 4; Mismatches 11; Indels 8; Gaps 2;

QY 1 PIG-----NL-GNNVNGNHLIPAPPLPSQTDGAA 29
 Db 93 PVIQISEPKTSNLCGNRAVYKSLIPVPRISVKTSASA 129

RESULT 37
 ADI40526
 ID ADI40526 standard; protein; 226 AA.

AC ADI40526;

DT 22-APR-2004 (first entry)

XX Human purified secretory polypeptide (SPTM), seq id 219.

XX Antiartherosclerotic; antipsoriatic; cytostatic; secretory molecule;
 KW agonist; antagonist; gene therapy; antisense; human; secretory;
 KW purified secretory polynucleotide; sptm; toxicity; arteriosclerosis;
 XX psoriasis; cancer.

OS Homo sapiens.

XX WO2003062385-A2.

PD 31-JUL-2003.

PF 15-JAN-2003; 2003WO-US001605.

XX 17-JAN-2002; 2002US-0349413P.

PR 17-JAN-2002; 2002US-0349946P.

XX (INCY-) INCYTE GENOMICS INC.

PA (JONE/) JONES A L.

PA (DAHL/) DAHL C R.

PA (GIET/) GIETZEN D.

PA (CHIN/) CHINN J.

PA (DUFO/) DUFOUR G E.

PA (JACK/) JACKSON J L.

PA (YUJY/) YU J Y.

PA (TUAS/) TUASON O.

PA (YAPP/) YAP P E.

PA (AMSH/) AMSHEY S R.

PA (DAWT/) DAW T C.

PA (LIUT/) LIU T F.
 PA (GERS/) GERSTIN E H.
 PA (PERA/) PERALTA C H.
 PA (LEWT/) LEWIS S A.
 PA (CHEN/) CHEN A J.
 PA (MARW/) MARWAHA R.
 PA (LANR/) LAN R Y.
 PA (URAS/) URASHKA M E.
 PA (KRIS/) KRISTNAM S R.
 PA (KOLL/) KOLLURU V.
 PA (PANE/) PANESAR I S.
 XX Jones AL, Dahl CR, Gietzen D, Chinn J, Dufour GE, Jackson JL,
 PI Yu JY, Tuason O, Yap PE, Ameshey SR, Dam TC, Liu TF, Gerstin EH,
 PI Peralta CH, Lewis SA, Chen AJ, Marwaha R, Lan RY, Urashka ME,
 PI Kristnam SR, Kolluru V, Panesar IS;
 DR WPI; 2003-853444/79.
 DR N-PSDB; ADI40363.

PT New isolated secreted polynucleotide for diagnosing or treating
 PT conditions, diseases or disorders associated with cell signaling e.g.
 PT arteriosclerosis, psoriasis, and cancer.

PS Claim 27a; SEQ ID NO 219; 486bp; English.

XX The invention relates to isolated purified secreted polynucleotides
 CC (sptm) (I), and the polypeptides (SPTM) encoded by sptm. A polypeptide
 CC encoded by (I) is used to identify a compound which binds to it. A
 CC microarray comprising (I) is used to generate a transcript. (I) is used
 CC to screen a compound for effectiveness in altering expression of (I). (I)
 CC is used to assess toxicity of a test compound. An agonist or antagonist
 CC identified by a new method is used in a pharmaceutical composition. The
 CC secretory molecules are used to diagnose or treat conditions, diseases or
 CC disorders associated with cell signaling e.g. arteriosclerosis,
 CC psoriasis, and cancer. Sequences given in ADI40308-ADI40468 represent
 CC human purified secretory polynucleotides of the invention, and those
 CC given in ADI40469-ADI40631 represent the polypeptides they encode.

XX Sequence 226 AA;

Query Match 32.7%; Score 54; DB 7; Length 226;
 Best Local Similarity 47.6%; Pred. No. 1.3e+02;
 Matches 10; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 10 NGNHLIPAPPLPSQTDGAAR 30
 Db 42 SGKLSPVPRPRPTQTASPAR 62

RESULT 38
 ADM06052

ID ADM06052 standard; protein; 279 AA.

AC ADM06052;

DT 20-MAY-2004 (first entry)

XX Human protein of the invention SEQ ID NO:4737.

XX human; gene therapy; diagnostic marker; pharmaceutical.

OS Homo sapiens.

XX EP1347046-A1.

XX 24-SEP-2003.

XX 12-APR-2002; 2002EP-00008400.

XX 22-MAR-2002; 2002JP-00137785.

XX (REAS-) RES ASSOC BIOTECHNOLOGY.

XX Iisogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;
XX WPI; 2003-723558/69.
DR N-PSDB; ADM03609.
XX
PT New polynucleotides and polypeptides are useful in gene therapy, for
PT developing a diagnostic marker or medicines for regulating their
PT expression and activity, or as a target of gene therapy.
XX
PS Claim 1; SEQ ID NO 4737; 305pp; English.
XX
CC The invention relates to a novel human polynucleotide and the encoded
CC polypeptide. A polynucleotide of the invention may have a use in gene
CC therapy. An oligonucleotide of the invention ADM06202-ADM06773 is useful
CC as a primer for synthesizing the polynucleotide or as a probe for
CC detecting the polynucleotide. The polynucleotides ADM01316-ADM03758 are
CC useful in gene therapy, for developing a diagnostic marker or medicines
CC for regulating their expression and activity, or as a target of gene
CC therapy. The proteins ADM03759-ADM06201 encoded by the polynucleotides
CC are useful as pharmaceutical agents. The present sequence represents a
CC protein sequence of the invention.
XX
SQ Sequence 279 AA;

Query Match 32.7%; Score 54; DB 7; Length 279;
Best Local Similarity 47.6%; Pred. No. 1.6e+02;
Matches 10; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

Oy 10 NGNHLIPAPPPLPSQTDGAA 30
: | : | | | | : | | | |
Db 95 SGKLSPPVPRPPTGTASPAR 115

RESULT 39
AEC88982
ID AEC88982 standard; protein; 279 AA.

AC AEC88982;

DT 01-DEC-2005 (first entry)

DE Human cDNA clone protein TKIDN20005210, SEQ ID 4737.

XX Osteopathic; Cytostatic; Antiinflammatory; Gastrointestinal Gen.;

KW Anticancer; Gene Therapy; Osteoporosis; Cancer; Inflammation; Gastritis;

XX stomach ulcer; gastrointestinal ulcer.

OS Homo sapiens.

XX EPI580263-A1.

XX 28-SEP-2005.

XX 12-APR-2002; 2004EP-00027348.

XX 22-MAR-2002; 2002JP-00137785.

XX 12-APR-2002; 2002EP-00008400.

XX (REAS-) RES ASSOC BIOTECHNOLOGY.

XX Iisogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;

XX Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;

XX Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;

XX WPI; 2005-667421/69.

XX N-PSDB; AEC86539.

XX New full-length cDNA sequences, useful for treating diseases, e.g.

XX osteoporosis, cancer, inflammation, gastritis, or gastroduodenal ulcer.

PS Example 3; SEQ ID NO 4737; 296pp; English.

XX The present invention relates to novel human cDNAs (AEC84246-AEC86688)
CC encoding proteins AEC86689-AEC89131. The cDNAs are useful for analyzing
CC the functions of the proteins, and for developing medicines for diseases
CC e.g. osteoporosis, cancer, inflammation, gastritis, or gastroduodenal
CC ulcer. Note: The sequence data for this patent did not form part of the
CC printed specification but was obtained in electronic format directly from
CC EPO.

XX Sequence 279 AA;

Query Match 32.7%; Score 54; DB 9; Length 279;
Best Local Similarity 47.6%; Pred. No. 1.6e+02;
Matches 10; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

Oy 10 NGNHLIPAPPPLPSQTDGAA 30
: | : | | | | : | | | |
Db 95 SGKLSPPVPRPPTGTASPAR 115

RESULT 40
ADY24521

ID ADY24521 standard; protein; 319 AA.

AC ADY24521;

DT 21-APR-2005 (first entry)

DE Plant full length insert polypeptide seqid 72305.

XX plant protectant; plant growth regulant; gene therapy; plant;

KW recombinant DNA construct; physical array; plant breeding marker;

KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;

KW extreme osmotic condition; pathogen tolerance; pest tolerance;

KW growth rate; cell cycle pathway; disease resistance;

KW galactomannan production; lignin production; plant growth regulator;

KW yield; plant growth; plant development; seed oil; protein yield;

XX protein content.

XX Unidentified.

XX US2004034888-A1.

XX 19-FEB-2004.

XX 28-APR-2003; 2003US-00425114.

XX 06-MAY-1999; 99US-00304517.

XX 05-NOV-2001; 2001US-00985678.

XX (LIUJ/) LIU J.

XX (ZHOU/) ZHOU Y.

XX (KOVA/) KOVALIC D K.

XX (SCRE/) SCREEN S E.

XX (TABAA/) TABASKA J E.

XX (CAOY/) CAO Y.

XX Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JB, Cao Y;

XX WPI; 2004-180133/17.

XX New recombinant DNA construct, useful for improving plant tolerance to

XX cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or

XX pests, for conferring increased resistance to plant disease, or for

XX improving yield.

XX Claim 1; SEQ ID NO 72305; 15pp; English.

XX The invention describes a recombinant DNA construct comprising a

XX polynucleotide consisting of a sequence encoding an amino acid sequence

XX available in electronic form from the US patent office at

XX ftp.segdata.uspto.gov/sequence.html?docid:2004034888. The polynucleotide

CC of the invention are also useful in physical arrays of molecules and as
CC plant breeding markers. The recombinant DNA construct is useful for
CC improving plant tolerance to cold, heat, drought, herbicides, extreme
CC osmotic conditions, pathogens or pests, for manipulating growth rate in
CC plant cells by modification of the cell cycle pathway, for conferring
CC increased resistance to plant disease, for producing galactomannan,
CC lignin or plant growth regulators, for increasing the rate of homologous
CC recombination in plants, for improving yield by modification of
CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
CC or by providing improved plant growth and development under at least one
CC stress condition or for modifying seed oil or protein yield and/or
CC content. This is the amino acid sequence of a plant full length insert
CC polypeptide that can be used in the recombinant DNA construct of the
CC invention.
XX

XX
SQ Sequence 319 AA;

Query Match 32.7%; Score 54; DB 8; Length 319;
Best Local Similarity 52.6%; Pred. No. 1.8e+02;

Matches 10; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 4 NLGNVNVGNHLIPPPPLP 22
|||||
:|||||

Db 124 NLGNMTPDEDOLPPPPPLP 142

Search completed: August 1, 2006, 21:41:38
Job time : 203 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 1, 2006, 21:38:25 ; Search time 298 Seconds
(without alignments)
93.122 Million cell updates/sec

Title: US-09-189-415D-7
Perfect score: 165
Sequence: 1 PIGNLGNVNGNHLIPAPPLPSQTDGAAR 30

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2849598 seqs, 925015592 residues
Total number of hits satisfying chosen parameters: 2849598

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : Uniprot_7.2.*
1: uniprot_prot.*
2: uniprot_crembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	165	100.0	549	2	O50190 ECOLI
2	165	100.0	550	2	O52147 ECOLI
3	165	100.0	550	2	O5WMC3 ECOLI
4	159	96.4	552	2	O9KWH9 ECOLI
5	149	90.3	551	2	O68258 ECOLI
6	149	90.3	551	2	O4ZIM0 ECOLI
7	149	90.3	551	2	O4ZIM1 ECOLI
8	132	80.0	547	2	O7BHL5 ECOLI
9	132	80.0	547	2	O7BHL5 ECOLI
10	132	80.0	547	2	O9FTI1 ECOLI
11	123	74.5	558	2	O85506 ECOLI
12	123	74.5	558	2	O4ZIM4 ECOLI
13	123	74.5	558	2	O4ZIM4 ECOLI
14	123	74.5	558	2	O9R396 ECOLI
15	123	74.5	558	2	O9R396 ECOLI
16	120	72.7	538	2	O58188 ECOLI
17	120	72.7	538	2	O58188 ECOLI
18	120	72.7	538	2	O47016 ECOLI
19	120	72.7	538	2	O47016 ECOLI
20	120	72.7	538	2	O58189 ECOLI
21	120	72.7	538	2	O58190 ECOLI
22	120	72.7	538	2	O58190 ECOLI
23	111	67.3	538	2	O58187 ECOLI
24	66.5	40.3	1554	2	O34G79 RHOPA
25	66.5	39.1	1510	2	O34G79 RHOPA
26	62	37.6	980	2	O8H3A7 ORYSA
27	61	37.0	279	2	O4FX35 LITMA
28	61	37.0	279	2	O86UN5 DICDI
29	60.5	36.7	116	2	O8HYJ8 FELCA
30	60.5	36.7	290	2	O8HYJ8 FELCA
31	60.5	36.7	589	2	O7TPJ8 RAT

32	60.5	36.7	712	2	O5K9X9 CRYNE
33	60.5	36.7	718	2	O55JM3 CRYNE
34	60.5	36.7	913	2	O5B437 EMENT
35	60	36.4	613	2	O9VKM2 DROME
36	60	36.4	938	2	O7YYA2 CRYPV
37	60	36.4	974	2	O726D4 DESVH
38	60	36.4	9529	2	O2M1P9 MAGSA
39	59	35.8	329	2	O2QM44 ORYSA
40	58.5	35.5	303	2	O4DJY5 TRYCR
41	58.5	35.5	416	2	O2S328 GSPHI
42	58.5	35.5	827	2	O6PJ85 HUMAN
43	58.5	35.5	1117	2	O41IT9 GIBZE
44	58.5	35.5	2486	2	O75179 HUMAN
45	58	35.2	336	2	O59L14 CANAL
46	58	35.2	336	2	O59L14 CANAL
47	58	35.2	634	2	O59LJ8 CANAL
48	58	35.2	959	2	O16987 CAEEL
49	58	35.2	1377	2	O54ND2 DICDI
50	58	35.2	1567	1	FMN2 MOUSE
51	58	35.2	1578	2	O505D3 MOUSE
52	57.5	34.8	331	2	O55BZ5 DICDI
53	57	34.5	228	2	O7Q266 ANOGA
54	57	34.5	241	2	O9VA32 DROME
55	57	34.5	273	2	O3SDG6 PARTE
56	57	34.5	365	2	O9XUP5 CAEEL
57	57	34.5	522	2	O9XEV3 ORYSA
58	57	34.5	696	2	O86S41 CAEEL
59	57	34.5	746	2	O6FVA0 CANGA
60	57	34.5	746	2	O67MJ9 SYMTH
61	57	34.5	760	2	O44760 CAEEL
62	57	34.5	833	2	O9YXET GYVRU
63	57	34.5	1729	2	O61A04 CAEEL
64	57	34.5	1752	2	O811K8 PLAVI
65	57	34.5	2148	2	O81P15 DROME
66	57	34.5	2206	2	O7PUP9 ANOGA
67	57	34.5	2625	2	O8MWZ9 DICDI
68	56.5	34.2	216	2	O5A0H9 EMENT
69	56.5	34.2	216	2	O5ICCA EMENT
70	56.5	34.2	487	2	O08989 YEAST
71	56.5	34.2	1224	2	O54DC8 DICDI
72	56	33.9	187	2	O2QP65 ORYSA
73	56	33.9	230	2	O4S158 TETNG
74	56	33.9	295	2	O8VBZ2 MSAV
75	56	33.9	295	2	O91LK7 MSAV
76	56	33.9	410	1	AROA PYRAB
77	56	33.9	551	2	O9HND4 PSEAE
78	56	33.9	938	2	O5CLY0 CRYHO
79	56	33.9	977	2	O4RPH3 USTMA
80	56	33.9	995	2	O4RPH3 USTMA
81	56	33.9	1249	2	O4G6R1 LITMA
82	56	33.9	1420	2	O13736 SCHPO
83	56	33.9	1459	2	O4PHQ3 USTMA
84	56	33.9	1738	2	O4PEFI USTMA
85	56	33.9	3402	2	O6S6P7 GALPH
86	56	33.9	3421	1	TBEGU EHYVB
87	55.5	33.6	127	2	O4DJ20 TRYCR
88	55.5	33.6	127	2	O4DM34 TRYCR
89	55.5	33.6	264	2	O2ULU5 ASPOR
90	55.5	33.6	315	2	O683P3 CVB
91	55.5	33.6	439	2	O6CB80 VARLI
92	55.5	33.6	624	2	O08912 YEAST
93	55.5	33.6	675	1	BCR ARDAE
94	55.5	33.6	1280	2	O903Y4 ABDAL
95	55.5	33.6	1732	2	O54GFS DICDI
96	55	33.3	237	2	O5AL52 CANAL
97	55	33.3	248	2	O3YVT3 SHISS
98	55	33.3	317	2	O4D7G8 TRYCR
99	55	33.3	347	2	O2ULI2 ASPOR
100	55	33.3	347	2	O874X7 PODAN

ALIGNMENTS

O5K9X9	CRYLOCOCU
O55JM3	CRYLOCOCU
O5B437	ASPERGILLUS
O9VKM2	DROSOPHILA
O7YYA2	CRYPTOSPORI
O726D4	DESULFOVIBR
O2M1P9	MAGNETOSPIR
O2QM44	ORYZA SATIV
O4DJY5	ORYZA SATIV
O2S328	SALEINIBACTE
O6PJ85	HOMO SAPIEN
O41IT9	GIBBERELLA
O75179	HOMO SAPIEN
O59L14	CANDIDA ALB
O59LJ8	CANDIDA ALB
O5FC83	CAENORHABDI
O16987	CAENORHABDI
O54ND2	DICTYOSTELI
O91104	MUS MUSCULU
O505D3	MUS MUSCULU
O55BZ5	DICTYOSTELI
O7Q266	ANOPHELES G
O9VA32	DROSOPHILA
O3SDG6	PARAMECIUM
O9XUP5	CAENORHABDI
O9XEV3	ORYZA SATIV
O86S41	CAENORHABDI
O6FVA0	CANDIDA GLA
O67MJ9	SYMBIOPACTE
O44760	CAENORHABDI
O9YXET	SPHAROPSIDA
O61A04	CAENORHABDI
O811K8	PLASMODIUM
O81P15	DROSOPHILA
O7PUP9	ANOPHELES G
O8MWZ9	DICTYOSTELI
O5A0H9	ASPERGILLUS
O5ICCA	EMERICELLA
O08989	SACCHAROMYC
O54DC8	DICTYOSTELI
O2QP65	ORYZA SATIV
O4S158	TETRAODON N
O8VBZ2	WHITE SPOT
O91LK7	WHITE SPOT
O9HND4	PSAUDOMONAS
O5CLY0	CRYPTOSPORI
O4RPH3	USULLAGO MA
O4PEFI	USULLAGO MA
O6S6P7	EQUID HERPE
TBEGU	EQUID HERPE
O4DJ20	TRYPANOSOMA
O4DM34	TRYPANOSOMA
O2ULU5	ASPERGILLUS
O683P3	CHRYSANTHEM
O6CB80	YARROWIA LI
O08912	SACCHAROMYC
BCR	AEDES AEGYP
O903Y4	AEDES ALBO
O54GFS	DICTYOSTELI
O5AL52	CANDIDA ALB
O3YVT3	SHIGELLA SO
O4D7G8	TRYPANOSOMA
O2ULI2	ASPERGILLUS
O874X7	PODOSPOIRA A

```
RESULT 1
OS0190_ECOLI PRELIMINARY; PRT; 549 AA.
ID 050190_ECOLI
AC 050190;
DT 01-JUN-1998, integrated into UniProtKB/TrEMBL.
DT 01-JUN-1998, sequence version 1.
DT 07-FEB-2006, entry version 19.
DE Translocated intimin receptor (Fragment).
GN Name=tir;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=E2348/69;
RX MEDLINE=98050926; PubMed=9390560; DOI=10.1016/S0092-8674(00)80437-7;
RA Kenny B., Devlinney R., Stein M., Reinscheid D.J., Frey E.A.,
RA Finlay B.B.;
RA "Enteropathogenic E. coli (EPEC) transfers its receptor for intimate
adherence into mammalian cells.";
RT Cell 91:511-520(1997).
RL [2]
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=E2348/69;
RA Stein M.S., Kenny B., Finlay B.B.;
RA Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs license
CC -----
EMBL: AF013122; AAB88410.1; -; Genomic_DNA.
DR HSBP; O9KWH9; 1F02.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR003536; TIR rcpt.
DR Pfam; PF07489; TIR_receptor_C; 1.
DR Pfam; PF03549; TIR_receptor_M; 1.
DR Pfam; PF07490; TIR_receptor_N; 1.
DR PRINTS; PRO1370; TRANSINTIMINR.
KW Receptor.
KW NON_TER.
SQ SEQUENCE 549 AA; 56843 MW; 40C8B8B234409A08 CRC64;

Query Match 100.0%; Score 165; DB 2; Length 549;
Best Local Similarity 100.0%; Pred. No. 5,3e-12;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PIGNLGNVNVGNHLLPPAPPLPSQTDGAAR 30
DB 2 PIGNLGNVNVGNHLLPPAPPLPSQTDGAAR 31

RESULT 2
OS2147_ECOLI PRELIMINARY; PRT; 550 AA.
ID 052147_ECOLI
AC 052147;
DT 01-JUN-1998, integrated into UniProtKB/TrEMBL.
DT 01-JUN-1998, sequence version 1.
DT 07-FEB-2006, entry version 18.
DE tir.
GN Name=tir;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=E2348/69;
RX MEDLINE=98254123; PubMed=9593291;
RA Elliott S.J., Wainwright L.A., McDaniel T.K., Jarvis K.G., Deng Y.K.,
RA Lai L.C., McNamara B.P., Sonnenberg M.S., Kaper J.B.,
```

```
RT "The complete sequence of the locus of enterocyte effacement (LEE)
RT from enteropathogenic Escherichia coli E2348/69.",
RT Mol. Microbiol. 28:1-4(1998).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs license
CC -----
EMBL: AF022236; AAC38390.1; -; Genomic_DNA.
DR HSBP; O9KWH9; 1F02.
DR SMR; OS2147; 272-336.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR003536; TIR rcpt.
DR Pfam; PF07489; TIR_receptor_C; 1.
DR Pfam; PF03549; TIR_receptor_M; 1.
DR Pfam; PF07490; TIR_receptor_N; 1.
DR PRINTS; PRO1370; TRANSINTIMINR.
SQ SEQUENCE 550 AA; 56510 MW; 19DD08A9B8251CB CRC64;

Query Match 100.0%; Score 165; DB 2; Length 550;
Best Local Similarity 100.0%; Pred. No. 5,3e-12;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PIGNLGNVNVGNHLLPPAPPLPSQTDGAAR 30
DB 2 PIGNLGNVNVGNHLLPPAPPLPSQTDGAAR 31

RESULT 3
OSWMC9_ECOLI PRELIMINARY; PRT; 550 AA.
ID 05WMC9_ECOLI
AC 05WMC9;
DT 23-NOV-2004, integrated into UniProtKB/TrEMBL.
DT 23-NOV-2004, sequence version 1.
DT 07-FEB-2006, entry version 7.
DE tir, translocated intimin receptor.
GN Name=tir;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=0181-6/86;
RX PubMed=15501811; DOI=10.1128/IAI.72.11.6722-6728.2004;
RA Gartner J.F., Schmidt M.A.;
RA "Comparative Analysis of Locus of Enterocyte Effacement Pathogenicity
RT Islands of Atypical Enteropathogenic Escherichia coli.",
RT Infect. Immun. 72:6722-6728(2004).
RL [2]
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=0181-6/86;
RA Gaertner J.;
RA "Etablierung eines Zellkulturmodells fuer M-Zellen und vergleichende
RT molekulare Analyse der LEE-Pathogenitätsinsel atypischer
RT enteropathogener Escherichia coli.";
RL Thesis (2002), Department of Biological Sciences, University of
Muenster, Muenster, Germany.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs license
CC -----
EMBL: AJ633129; CAG17536.1; -; Genomic_DNA.
DR SMR; OSWMC9; 272-336.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR003536; TIR rcpt.
DR Pfam; PF07489; TIR_receptor_C; 1.
DR Pfam; PF03549; TIR_receptor_M; 1.
DR Pfam; PF07490; TIR_receptor_N; 1.
DR PRINTS; PRO1370; TRANSINTIMINR.
KW Receptor.
SQ SEQUENCE 550 AA; 56499 MW; 126029DC070E9BB CRC64;
```

Query Match 100.0%; Score 165; DB 2; Length 550;
Best Local Similarity 100.0%; Pred. No. 5, 3e-12;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PIGNLGNVNGNHLPPAPPLPSQTDGAAR 30
DB 2 PIGNLGNVNGNHLPPAPPLPSQTDGAAR 31

RESULT 4

Q9KMH9_ECOLI PRELIMINARY; PRT; 552 AA.
AC Q9KMH9;
DT 01-OCT-2000, integrated into UniProtKB/TrEMBL.
DT 01-OCT-2000, sequence version 1.
DE 07-FEB-2006, entry version 13.
DE Tlr.
GN Name=tlr;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=HK01;
RA Abe A., Nagano H.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs license

DR EMBL; AB036053; BA96815.1; -; Genomic_DNA.
DR PDB; 1P02; X-ray; T=272-336.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR003536; Tlr rcpt.
DR Pfam; PF07489; Tlr_receptor_C; 1.
DR Pfam; PF03549; Tlr_receptor_M; 1.
DR Pfam; PF07490; Tlr_receptor_N; 1.
DR PRINTS; PR01370; TRANSINTIMINR.
SQ SEQUENCE 552 AA; 57005 MW; ABD79EB22E50A4B CRC64;

Query Match 96.4%; Score 159; DB 2; Length 552;
Best Local Similarity 96.7%; Pred. No. 3e-11;
Matches 29; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PIGNLGNVNGNHLPPAPPLPSQTDGAAR 30
DB 2 PIGNLGNVNGNHLPPAPPLPSQTDGAAR 31

RESULT 5

O68258_ECOLI PRELIMINARY; PRT; 551 AA.
AC O68258;
DT 01-AUG-1998, integrated into UniProtKB/TrEMBL.
DT 01-AUG-1998, sequence version 1.
DE 07-FEB-2006, entry version 19.
DE Translocated intimin receptor.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN NUCLEOTIDE SEQUENCE.
RP MEDLINE=98187918; PubMed=9529069;
RA Voss E., Paton A.W., Manning P.A., Paton J.C.;
RT "Molecular analysis of Shiga toxin-producing Escherichia coli O111:H-
proteins which react with sera from patients with hemolytic-uremic
syndrome.";
RL Infect. Immun. 66:1467-1472(1998).
RN [2]

RP NUCLEOTIDE SEQUENCE.

RX MEDLINE=99003184; PubMed=9784578;
RA Paton A.W., Manning P.A., Woodrow M.C., Paton J.C.;
RT "Translocated intimin receptor (Tlr) of Shiga-toxinogenic Escherichia
coli isolates belonging to serogroups O26, O111, and O157 react with
sera from patients with hemolytic-uremic syndrome and exhibit marked
sequence heterogeneity.";
RT Infect. Immun. 66:5580-5586(1998).
RL
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs license

DR EMBL; AF025311; AAC69249.1; -; Genomic_DNA.
DR HSSP; Q9KMH9; 1P02.
DR SMR; O68258; 272-336.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR003536; Tlr rcpt.
DR Pfam; PF07489; Tlr_receptor_C; 1.
DR Pfam; PF03549; Tlr_receptor_M; 1.
DR Pfam; PF07490; Tlr_receptor_N; 1.
DR PRINTS; PR01370; TRANSINTIMINR.
SQ SEQUENCE 551 AA; 56975 MW; 6EC95F76F0F44CC CRC64;

Query Match 90.3%; Score 149; DB 2; Length 551;
Best Local Similarity 90.0%; Pred. No. 5, 3e-10;
Matches 27; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 PIGNLGNVNGNHLPPAPPLPSQTDGAAR 30
DB 2 PIGNLGNVNGNHLPPAPPLPSQTDGAAR 31

RESULT 6

O42IM0_ECOLI PRELIMINARY; PRT; 551 AA.
AC O42IM0;
DT 07-JUN-2005, integrated into UniProtKB/TrEMBL.
DT 07-JUN-2005, sequence version 1.
DT 07-FEB-2006, entry version 5.
DE Translocated intimin receptor.
GN Name=tlr;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=CPG124-MC416, and CPG123-G58;
RX PubMed=16272509; DOI=10.1128/JCM.43.11.5715-5720.2005;
RA Garmendia J., Ren Z., Tennant S., Midoll Viera M.A., Chong Y.,
RA Whale A., Azopardo K., Dahan S., Sicilli M.P., Franzolin M.R.,
RA Trabulsi L.R., Phillips A., Gomes T.A., Xu J., Robins-Browne R.,
RA Frankel G.;
RT "Distribution of tccp in clinical Enterohemorrhagic and
RT Enteropathogenic Escherichia coli isolates.";
RL J. Clin. Microbiol. 43:5715-5720(2005).
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs license

DR EMBL; DQ007024; AAY25395.1; -; Genomic_DNA.
DR EMBL; DQ007022; AAY25393.1; -; Genomic_DNA.
DR SMR; O42IM0; 272-336.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR003536; Tlr rcpt.
DR Pfam; PF07489; Tlr_receptor_C; 1.
DR Pfam; PF03549; Tlr_receptor_M; 1.
DR Pfam; PF07490; Tlr_receptor_N; 1.
DR PRINTS; PR01370; TRANSINTIMINR.
KM Receptor.

SO	SEQUENCE	551 AA;	57066 MW;	83B899E6183AFB4	CRC64;
	Query Match	90.3%;	Score 149;	DB 2;	Length 551;
	Best Local Similarity	90.0%;	Pred. No. 5.3e-10;		
	Matches	27;	Conservative	2;	Mismatches 1;
Oy		1	PIGNLGNVNVNGNHLIPAPAPLPSPQTDGASR	30	
Db		2	PIGNLGNVNVNSNNLIPAPAPLPSPQTDGASR	31	
	RESULT 7				
	O4ZIM1_ECOLI		PRELIMINARY;	PRT;	551 AA.
ID	O4ZIM1_ECOLI				
AC	O4ZIM1_1				
OT	07-JUN-2005,		integrated into UniProtKB/TrEMBL.		
DT	07-JUN-2005,		sequence version 1.		
DT	07-FEB-2006,		entry version 4.		
DE	Translocated intimin receptor.				
GN	Name=tir;				
OS	Escherichia coli.				
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;				
OC	Enterobacteriaceae; Escherichia.				
OK	NCBI_TaxID=562;				
RP	[1]				
RN	NUCLEOTIDE SEQUENCE.				
RC	STRAIN=CPG421;				
RX	PubMed=16972509; DOI=10.1128/JCM.43.11.5715-5720.2005.				
RA	Garmendia J., Ren Z., Tennant S., Midollini Vierra M.A., Chong Y.,				
RA	Whale A., Azopardo K., Dahan S., Sircelli M.P., Franzolin M.R.,				
RA	Tribulsi L.R., Phillips A., Gomes T.A., Xu J., Robins-Browne R.,				
RA	Frankel G.;				
RT	"Distribution of tccp in Clinical Enterohemorrhagic and				
RL	Enteropathogenic Escherichia coli Isolates."				
RT	J. Clin. Microbiol. 43:5715-5720(2005).				
CC	-----				
CC	Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms				
CC	Distributed under the Creative Commons Attribution-NoDerivs License				
CC	-----				
DR	EMBL; DQ007023; AAY25394.1; -; Genomic_DNA.				
DR	SMR; O4ZIM1; 272-336.				
DR	GO; GO:0004872; F:receptor activity; IEA.				
DR	GO; GO:0007155; P:cell adhesion; IEA.				
DR	InterPro; IPR003536; Tir_tccp.				
DR	Pfam; PF07489; Tir_receptor_C; 1.				
DR	Pfam; PF03549; Tir_receptor_M; 1.				
DR	Pfam; PF07490; Tir_receptor_N; 1.				
DR	PRINTS; PR01370; TRANSINTIMINR.				
KW	Receptor.				
SO	SEQUENCE	551 AA;	57066 MW;	333DB01FC9B461CB	CRC64;
	Query Match	90.3%;	Score 149;	DB 2;	Length 551;
	Best Local Similarity	90.0%;	Pred. No. 5.3e-10;		
	Matches	27;	Conservative	2;	Mismatches 1;
Oy		1	PIGNLGNVNVNGNHLIPAPAPLPSPQTDGASR	30	
Db		2	PIGNLGNVNVNSNNLIPAPAPLPSPQTDGASR	31	
	RESULT 8				
ID	O7BHL5_9ENTR		PRELIMINARY;	PRT;	547 AA.
AC	O7BHL5_1				
DT	05-JUL-2004,		integrated into UniProtKB/TrEMBL.		
DT	05-JUL-2004,		sequence version 1.		
DT	07-FEB-2006,		entry version 8.		
DE	Translocated intimin receptor Tir.				
OS	Citrobacter rodentium.				
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;				
OC	Enterobacteriaceae; Citrobacter.				
OK	NCBI_TaxID=67825;				
RP	[1]				

```

RP NUCLEOTIDE SEQUENCE.
RC STRAIN=DSB1007;
RX MEDLINE=21437640; PubMed=11553577;
RA DOI=10.1128/JAI.69.10.6323-6335.2001;
RT Deng W., Li Y., Vallance B.A., Finlay B.B.;
RT "Locus of enterocyte effacement from Citrobacter rodentium: sequence
RT analysis and evidence for horizontal transfer among attaching and
RT effecting pathogens.";
RL Infect. Immun. 69:6323-6335 (2001).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs license
CC -----
DR EMBL: AF311901; AAL06376.1; -; Genomic_DNA.
DR SMR: Q7BH5T; 270-334.
DR GO: GO:0004872; F:receptor activity; IEA.
DR GO: GO:0007155; P:cell adhesion; IEA.
DR InterPro: IPR003536; TIR rcpt.
DR Pfam: PF07489; TIR_receptor_C_1.
DR Pfam: PF03549; TIR_receptor_M_1.
DR Pfam: PF07490; TIR_receptor_N_1.
DR PRINTS: PR01370; TRANSINTIMINR.
KW Receptor.
SQ SEQUENCE 547 AA; 56284 MW; 02CAC6D52FA6EBE1 CRC64;

Query Match 80.0%; Score 132; DB 2; Length 547;
Best Local Similarity 83.3%; Pred. No. 7,1e-08;
Matches 25; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 PIGNLNNVNGNHLIPAPLPSPQTGGAR 30
DB 2 PIGNLNNNNISNNLIIPAPLPSPQTDGATR 31
||||||| : |||||
||||||| : |||||

RESULT 9
ID O9ETI1_GENTR PRELIMINARY; PRT; 547 AA.
AC O9ETI1;
DT 01-MAR-2001, integrated into UniProtKB/TrEMBL.
DT 01-MAR-2001, sequence version 1.
DT 07-FEB-2006, entry version 12.
DE Translocated intimin receptor Tir.
OS Citrobacter rodentium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacterales;
OC Enterobacteriaceae; Citrobacter.
OX NCBI_TaxID=67825;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CDC 1843-73T and DSB100;
RX MEDLINE=20553330; PubMed=1101562;
RA Luperchio S.A., Newman J.V., Dangler C.A., Schenzel M.D.,
RA Brenner D.J., Steigertwald A.G., Schauer D.B.;
RT "Citrobacter rodentium, the causative agent of transmissible murine
RT colonic hyperplasia, exhibits clonality: synonymy of C. rodentium and
RT mouse-pathogenic Escherichia coli.";
RL J. Clin. Microbiol. 38:4343-4350(2000).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs license
CC -----
DR EMBL: AF301618; AAG40758.1; -; Genomic DNA.
DR EMBL: AF301617; AAG25642.1; -; Genomic_DNA.
DR HSSP: O9KMH9; 1F02.
DR SMR: O9ETI1; 270-334.
DR GO: GO:0004872; F:receptor activity; IEA.
DR GO: GO:0007155; P:cell adhesion; IEA.
DR InterPro: IPR003536; TIR rcpt.
DR Pfam: PF07489; TIR_receptor_C_1.
DR Pfam: PF03549; TIR_receptor_M_1.
DR Pfam: PF07490; TIR_receptor_N_1.
DR PRINTS: PR01370; TRANSINTIMINR.
KW Receptor.
SQ SEQUENCE 547 AA; 56271 MW; CB831B301049C37 CRC64;

```

Query Match 80.0%; Score 132; DB 2; Length 547;
Best Local Similarity 83.3%; Pred. No. 7.1e-08;
Matches 25; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 PIGNLGNVNGNHLIPAPPLPSQTDGAA 30
DB 2 PIGNLGNVNNISNNLIPAPPLPSQTDGATR 31

RESULT 10

Q9WXK1_ECOLI PRELIMINARY; PRT; 547 AA.
AC Q9WXK1;
DT 01-NOV-1999, integrated into UniProtKB/TrEMBL.
DT 01-NOV-1999, sequence version 1.
DE 07-FEB-2006, entry version 16.
DE Translocated intimin receptor.
GN Name=tir;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MPEC;
RA Okutani A., Itoh K., Sasakawa C.;
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.

CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NonDerivs license

DR EMBL: AB026719; BAA77400.1; -; Genomic_DNA.
DR HSSP: Q9KMH9; 1F02.
DR SMR; Q9WXK1; 270-334.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro: IPR003536; TIR rcpt.
DR Pfam: PR07489; TIR_receptor_C; 1.
DR Pfam: PF03549; TIR_receptor_M; 1.
DR Pfam: PF07490; TIR_receptor_N; 1.
DR PRINTS: PR01370; TRANSINTIMINR.
DR Receptor.
SQ SEQUENCE 547 AA; 56284 MW; 02CAC6D625FA6EB1 CRC64;

Query Match 80.0%; Score 132; DB 2; Length 547;
Best Local Similarity 83.3%; Pred. No. 7.1e-08;
Matches 25; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 PIGNLGNVNGNHLIPAPPLPSQTDGAA 30
DB 2 PIGNLGNVNNISNNLIPAPPLPSQTDGATR 31

RESULT 11

O85506_ECOLI PRELIMINARY; PRT; 558 AA.
AC O85506;
DT 01-NOV-1998, integrated into UniProtKB/TrEMBL.
DT 01-NOV-1998, sequence version 1.
DT 07-FEB-2006, entry version 18.
DE Translocated intimin receptor Tir.
GN Name=tir;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=955f2;
RC MEDLINE=99003184; PubMed=9784578;
RA Paton A.W., Manning P.A., Woodrow M.C., Paton J.C.;
RT "Translocated intimin receptors (Tir) of Shiga-toxinogenic Escherichia

RT coli isolates belonging to serogroups O26, O111, and O157 react with
RT sera from patients with hemolytic-uremic syndrome and exhibit marked
RT sequence heterogeneity.";
RL Infect. Immun. 66:5580-5586(1998).

CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NonDerivs license

DR EMBL: AF070067; AAC69314.1; -; Genomic_DNA.
DR HSSP; Q9KMH9; 1F02.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro: IPR003536; TIR rcpt.
DR Pfam: PR07489; TIR_receptor_C; 1.
DR Pfam: PF03549; TIR_receptor_M; 1.
DR Pfam: PF07490; TIR_receptor_N; 1.
DR PRINTS: PR01370; TRANSINTIMINR.
DR Receptor.
KW SEQUENCE 558 AA; 58176 MW; CA2CDAC94527C2E CRC64;

Query Match 74.5%; Score 123; DB 2; Length 558;
Best Local Similarity 82.1%; Pred. No. 9.7e-07;
Matches 23; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 PIGNLGNVNGNHLIPAPPLPSQTDGA 28
DB 2 PIGNLGNVNNISNNLIPAPPLPSQTDGA 29

RESULT 12

O4ZIM4_ECOLI PRELIMINARY; PRT; 558 AA.
AC O4ZIM4;
DT 07-JUN-2005, integrated into UniProtKB/TrEMBL.
DT 07-JUN-2005, sequence version 1.
DT 07-FEB-2006, entry version 5.
DE Translocated intimin receptor.
GN Name=tir;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CPG122-657, and CPG6;
RC PubMed=16272509; DOI=10.1128/JCM.43.11.5715-5720.2005;
RA Garmendia J., Ren Z., Tennant S., Midollil Viera M.A., Chong Y.,
RA Whale A., Azopardo K., Dahan S., Siccoli M.P., Franzolin M.R.,
RA Trubelsi L.R., Phillips A., Gomes T.A., Xu J., Robins-Browne R.,
RA Frankel G.;
RT "Distribution of tccp in Clinical Enterohemorrhagic and
RT Enteropathogenic Escherichia coli Isolates";
RL J. Clin. Microbiol. 43:5715-5720(2005).

CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NonDerivs license

DR EMBL: DQ007020; AAY25391.1; -; Genomic_DNA.
DR EMBL: DQ007019; AAY25390.1; -; Genomic_DNA.
DR SMR; O4ZIM4; 271-335.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro: IPR003536; TIR rcpt.
DR Pfam: PR07489; TIR_receptor_C; 1.
DR Pfam: PF03549; TIR_receptor_M; 1.
DR Pfam: PF07490; TIR_receptor_N; 1.
DR PRINTS: PR01370; TRANSINTIMINR.
DR Receptor.
KW SEQUENCE 558 AA; 58008 MW; 69E1ADF5E80AFB10 CRC64;

Query Match 74.5%; Score 123; DB 2; Length 558;
Best Local Similarity 82.1%; Pred. No. 9.7e-07;
Matches 23; Conservative 2; Mismatches 3; Indels 0; Gaps 0;


```
AC Q58188;
DT 26-APR-2005, integrated into UniProtKB/TrEMBL.
DT 26-APR-2005, sequence version 1.
DT 07-FEB-2006, entry version 6.
DE Translocated intimin receptor.
GN Name=tit;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Baileman's Run stream;
RX PubMed=15870341; DOI=10.1128/AEM.71.5.2511-2519.2005;
RA Higgins J.A., Belt K.T., Karns J.S., Russell-Anelli J., Shelton D.R.;
RT "titr- and sex-positive Escherichia coli in Stream Waters in a
RT Metropolitan Area.";
RL Appl. Environ. Microbiol. 71:2511-2519(2005).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Baileman's Run stream;
RA Hohn C., Karns J.S., Higgins J.A.;
RL Submitted (FEB-2005) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs license
-----
DR EMBL, AY944737, AAX47730.1; -, Genomic DNA.
DR SMR, Q58188, 287-351.
DR GO, GO:0004872, F:receptor activity; IEA.
DR GO, GO:0007155, P:cell adhesion; IEA.
DR InterPro, IPR003536; Titr rcpt.
DR Pfam, PF07489; Titr_receptor_C; 1.
DR Pfam, PF03549; Titr_receptor_M; 1.
DR Pfam, PF07490; Titr_receptor_N; 1.
DR PRINTS, PR01370; TRANSINTIMINR.
DR Receptor.
KM
SQ SEQUENCE 574 AA; 59314 MW; DC78996D1BA3F5F0 CRC64;

Query Match 74.5%; Score 123; DB 2; Length 574;
Best Local Similarity 82.1%; Pred. No. 1e-06;
Matches 23; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 PIGNLGNVNGNHLIPAPPLPSQTDGA 28
Db 2 PIGNLGNVNVNVALIPAPPLPSQTDGA 29

RESULT 16
085508 ECOLI PRELIMINARY; PRT; 538 AA.
AC 085508;
DT 01-NOV-1998, integrated into UniProtKB/TrEMBL.
DT 01-NOV-1998, sequence version 1.
DT 07-FEB-2006, entry version 16.
DE Translocated intimin receptor Titr.
GN Name=tit;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=EPEC87A;
RX MEDLINE=99003184; PubMed=9784578;
RA Paton A.W., Manning P.A., Woodrow M.C., Paton J.C.;
RT "Translocated intimin receptors (Titr) of Shiga-toxicogenic Escherichia
RT coli isolates belonging to serogroups O26, O111, and O157 react with
RT sera from patients with hemolytic-uremic syndrome and exhibit marked
RT sequence heterogeneity.";
RL Infect. Immun. 66:5580-5586(1998).
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC
```

```
CC Distributed under the Creative Commons Attribution-NonDerivs license
-----
DR EMBL, AF070069; AAC69318.1; -, Genomic DNA.
DR HSSP, Q9KWH9, 1P02.
DR SMR, Q85508, 261-325.
DR GO, GO:0004872, F:receptor activity; IEA.
DR GO, GO:0007155, P:cell adhesion; IEA.
DR InterPro, IPR003536; Titr rcpt.
DR Pfam, PF07489; Titr_receptor_C; 1.
DR Pfam, PF03549; Titr_receptor_M; 1.
DR Pfam, PF07490; Titr_receptor_N; 1.
DR PRINTS, PR01370; TRANSINTIMINR.
DR Receptor.
KM
SQ SEQUENCE 538 AA; 55602 MW; 447052A0E3214D6D CRC64;

Query Match 72.7%; Score 120; DB 2; Length 538;
Best Local Similarity 82.1%; Pred. No. 2.2e-06;
Matches 23; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 PIGNLGNVNGNHLIPAPPLPSQTDGA 28
Db 2 PIGNLGNVNVNVALIPAPPLPSQTDGA 29

RESULT 17
Q47014 ECOLI PRELIMINARY; PRT; 538 AA.
AC Q47014;
DT 01-NOV-1996, integrated into UniProtKB/TrEMBL.
DT 01-MAY-1999, sequence version 2.
DT 07-FEB-2006, entry version 24.
DE Translocated intimin receptor Titr (Translocated intimin co-receptor)
DE (Espg protein).
GN Name=tit; Synonyms=espg;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=REPEC 84/110/L, and E65/56;
RA Krejany E.O.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=95ZG1;
RX MEDLINE=99003184; PubMed=9784578;
RA Paton A.W., Manning P.A., Woodrow M.C., Paton J.C.;
RT "Translocated intimin receptors (Titr) of Shiga-toxicogenic Escherichia
RT coli isolates belonging to serogroups O26, O111, and O157 react with
RT sera from patients with hemolytic-uremic syndrome and exhibit marked
RT sequence heterogeneity.";
RL Infect. Immun. 66:5580-5586(1998).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=B10;
RX MEDLINE=20187493; PubMed=10722617;
RX DOI=10.1128/IAI.68.4.2171-2182.2000;
RA Marches O., Neugayrede J.-P., Bouillier S., Mainil J., Charlier G.,
RA Raymond I., Pohl P., Bourry M., De Rycke J., Milton A., Oswald E.;
RT "Role of titr and intimin in the virulence of rabbit enteropathogenic
RT Escherichia coli serotype O103:H2.";
RL Infect. Immun. 68:2171-2182(2000).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=413/89-1;
RX MEDLINE=98294040; PubMed=9632251;
RA Deibel C., Kremer S., Chakraborty T., Ebel F.;
RT "Espg, a novel secreted protein of attaching and effacing bacteria, is
RT directly translocated into infected host cells where it appears as a
RT tyrosine-phosphorylated 90 kDa protein.";
RL Mol. Microbiol. 28:463-474(1998).
RN [5]
```

```

RP NUCLEOTIDE SEQUENCE.
RC STRAIN=413/89-1;
RA Benkel P., Chakraborty T.;
RL Submitted (Apr-2000) to the EMBL/genbank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
DR EMBL; U59502; AAC32028.2; -; Genomic DNA.
DR EMBL; AF070068; AAC69316.1; -; Genomic DNA.
DR EMBL; AF113597; AAP03080.1; -; Genomic DNA.
DR EMBL; AJ223063; CA11065.1; -; Genomic DNA.
DR EMBL; AJ277443; CAC81869.1; -; Genomic DNA.
DR EMBL; AF132728; AAD27868.1; -; Genomic DNA.
DR HSSP; O9KWH9; 1P02.
DR SMR; Q47014; 261-325.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR003536; T1r_rcpt.
DR Pfam; PF07489; T1r_receptor_C; 1.
DR Pfam; PF03549; T1r_receptor_M; 1.
DR Pfam; PF07490; T1r_receptor_N; 1.
DR PRINTS; PR01370; TRANSINTIMINR.
KW Receptor.
SQ SEQUENCE 538 AA; 55421 MW; 31D7A8E227B3D06C CRC64;

Query Match 72.7%; Score 120; DB 2; Length 538;
Best Local Similarity 82.1%; Pred. No. 2.2e-06;
Matches 23; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Oy 1 PIGNGNVNGNHLIPAPPLPSQTDGA 28
    |||||:| | ||||| |||||
Db 2 PIGNGNPNVVALIPAPPLPSQTDGA 29

RESULT 18
O47016_ECOLI PRELIMINARY; PRT; 538 AA.
AC Q47016;
DT 01-NOV-1996, integrated into UniProtKB/TrEMBL.
DT 01-AUG-1999, sequence version 2.
DT 07-FEB-2006, entry version 23.
DE Translocated intimin receptor (T1r).
GN Name=t1r;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OK NCBI_TaxID=562;
RN NCBI [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=REPEC 83/39;
RA O'Brien R.A.;
RL Submitted (DEC-1996) to the EMBL/genbank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=REPEC 83/39;
RA Krejany E.O.;
RL Submitted (MAR-1999) to the EMBL/genbank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=RDEC-1;
RA Agin T.S., Boedeker E.C.;
RL Submitted (JAN-1998) to the EMBL/genbank/DBJ databases.
RN [4]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=RDEC-1;
RA MEDLINE=96254123; PubMed=9593291;
RX Elliott S.J., Mainwright L.A., McDaniel T.K., Jarvis K.G., Deng Y.K.,
RA Lai L.C., McNamara B.P., Donnenberg M.S., Kaper J.B.;
RT "The complete sequence of the locus of enterocyte effacement (LEE)
RT from enteropathogenic Escherichia coli E2348/69.";
RL Mol. Microbiol. 28:1-4(1998).
RN [5]

```

```

RP NUCLEOTIDE SEQUENCE.
RC STRAIN=RDEC-1;
RX MEDLINE=97055784; PubMed=8900070; DOI=10.1016/0378-1097(96)00371-0;
RA Agin T.S., Catey J.R., Boedeker E.C., Wolf M.K.;
RT "Characterization of the eaeA gene from rabbit enteropathogenic
RT Escherichia coli strain RDEC-1 and comparison to other eaeA genes from
RT bacteria that cause attaching-effacing lesions.";
RL FEMS Microbiol. Lett. 144:249-258(1996).
RN [6]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=RDEC-1;
RX MEDLINE=21153569; PubMed=11254564;
RX DOI=10.1128/IAI.69.4.2107-2115.2001;
RA Zhu C., Agin T.S., Elliott S.J., Johnson L.A., Thate T.E., Kaper J.B.,
RA Boedeker E.C.;
RT "Complete nucleotide sequence and analysis of the locus of enterocyte
RT effacement from rabbit diarrheagenic Escherichia coli RDEC-1.";
RL Infect. Immun. 69:2107-2115(2001).
RN [7]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=83/39;
RX MEDLINE=22063667; PubMed=12067342;
RA Tauschek M., Strugnell R.A., Robins-Browne R.M.;
RT "Characterization and evidence of mobilization of the LEE
RT pathogenicity island of rabbit-specific strains of enteropathogenic
RT Escherichia coli.";
RL Mol. Microbiol. 44:1533-1550(2002).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
DR EMBL; U59504; AAD19750.1; -; Genomic DNA.
DR EMBL; AF045568; AAC15683.1; -; Genomic DNA.
DR EMBL; AF200363; AAK6722.1; -; Genomic DNA.
DR EMBL; AF453441; AAL57549.1; -; Genomic DNA.
DR HSSP; O9KWH9; 1P02.
DR SMR; Q47016; 261-325.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR003536; T1r_rcpt.
DR Pfam; PF07489; T1r_receptor_C; 1.
DR Pfam; PF03549; T1r_receptor_M; 1.
DR Pfam; PF07490; T1r_receptor_N; 1.
DR PRINTS; PR01370; TRANSINTIMINR.
KW Receptor.
SQ SEQUENCE 538 AA; 55411 MW; 3132A969B7B3D06C CRC64;

Query Match 72.7%; Score 120; DB 2; Length 538;
Best Local Similarity 82.1%; Pred. No. 2.2e-06;
Matches 23; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Oy 1 PIGNGNVNGNHLIPAPPLPSQTDGA 28
    |||||:| | ||||| |||||
Db 2 PIGNGNPNVVALIPAPPLPSQTDGA 29

RESULT 19
O58189_ECOLI PRELIMINARY; PRT; 538 AA.
AC Q58189;
DT 26-APR-2005, integrated into UniProtKB/TrEMBL.
DT 26-APR-2005, sequence version 1.
DT 07-FEB-2006, entry version 6.
DE Translocated intimin receptor.
GN Name=t1r;
OS Escherichia coli O157:H-.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OK NCBI_TaxID=183192;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Gwynns Run Gwynnbrook stream;
RX PubMed=15870341; DOI=10.1128/AEM.71.5.2511-2519.2005;

```



```

RA Higgins J.A., Belt K.T., Karns J.S., Russell-Anelli J., Shelton D.R.;
RT "tir- and stx-positive Escherichia coli in Stream Waters in a
RT Metropolitan Area.";
RL Appl. Environ. Microbiol. 71:2511-2519 (2005).
RN [2].
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Gwynns Run Gwynnbrook stream;
RA Hohn C., Shelton D.R., Higgins J.A.;
CC Submitted (FEB-2005) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
DR DR EMBL, AY944736; AAX4729.1; -; Genomic_DNA.
DR SMR, OS8189; 261-325.
DR GO, GO:0004872; F:receptor activity; IEA.
DR GO, GO:0007155; P:cell adhesion; IEA.
DR InterPro, IPR003536; TIR_rept.
DR Pfam, PF07489; TIR_receptor_C; 1.
DR Pfam, PF03549; TIR_receptor_M; 1.
DR Pfam, PF07490; TIR_receptor_N; 1.
DR PRINTS, PR01370; TRANSINTMINR.
KW Receptor.
KW SEQUENCE 538 AA; 55496 MW; 98848C2F508FC943 CRC64;

Query Match 72.7%; Score 120; DB 2; Length 538;
Best Local Similarity 82.1%; Pred. No. 2,2e-06;
Matches 23; Conservative 1; Mismatches 4; Indels 0; Gaps 0.

OY 1 PIGNLGNVNGNHLIPAPPLPSQTGDA 28
Db 2 PIGNLGNPNVRALIPAPPLPSQTGDA 29
|||||:|||||
|:|:|:|:|:|
|:|:|:|:|:|

RESULT 20
O58190_ECOLI PRELIMINARY; PRT; 538 AA.
ID O58190 ECOLI
AC O58190;
DT 26-APR-2005, integrated into UniProtKB/TrEMBL.
DT 26-APR-2005, sequence version 1.
DT 07-FEB-2006, entry version 6.
DE Translocated intimin receptor.
DE Name=tir;
GN Escherichia coli O157:H-.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OC NCBI_TaxID=183192;
CX [1]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=Spring Branch stream;
RX PubMed=15870341; DOI=10.1128/AEM.71.5.2511-2519.2005;
RA Higgins J.A., Belt K.T., Karns J.S., Russell-Anelli J., Shelton D.R.;
RT "tir- and stx-positive Escherichia coli in Stream Waters in a
RT Metropolitan Area.";
RT Appl. Environ. Microbiol. 71:2511-2519 (2005).
RN [2].
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Spring Branch stream;
RA Hohn C., Shelton D.R., Higgins J.A.;
CC Submitted (FEB-2005) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
DR DR EMBL, AY944735; AAX4728.1; -; Genomic_DNA.
DR SMR, OS8190; 261-325.
DR GO, GO:0004872; F:receptor activity; IEA.
DR GO, GO:0007155; P:cell adhesion; IEA.
DR InterPro, IPR003536; TIR_rept.
DR Pfam, PF07489; TIR_receptor_C; 1.
DR Pfam, PF03549; TIR_receptor_M; 1.
DR Pfam, PF07490; TIR_receptor_N; 1.
DR PRINTS, PR01370; TRANSINTMINR.

```

KM	Receptor.
SQ	SEQUENCE 538 AA; 55609 MW; 0336B5E18787C18E CRC64;
QY	Query Match 72.7%; Score 120; DB 2; Length 538; Best Local Similarity 82.1%; Pred. No. 2.2e-06; Matches 23; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
DB	1 PIGNLGNVNGNHLIPAPLPFSQTGDA 28 : 2 PIGNLGNPNVRALIPAPLPFSQTGDA 29
RESULT 21	
ID	OSKSP9_ECOLI PRELIMINARY; PRT; 538 AA.
AC	OSKSP9-
DT	15-FEB-2005, integrated into UniProtKB/TREMBL.
DT	07-FEB-2006, entry version 4.
DE	Translocated intimin receptor.
GN	Name=tir;
OS	Escherichia coli.
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC	Enterobacteriaceae; Escherichia.
OX	NCBI_Taxid=562;
[1]	
RN	NUCLEOTIDE SEQUENCE.
RC	STRAIN=RM1374;
RA	MEDLINE=21538660; PubMed=11682182;
RX	Jores J., Rumer L., Kiesel S., Kaper J.B., Wieler L.H.;
RT	"Identification of a new pathogenicity island inserted in the phev
RT	rRNA gene of the bovine Shiga toxin-producing E. coli strain RM1374
RT	(O103:H2) harboring a locus of enterocyte effacement that is flanked
RT	by intact insertion elements.";
RL	FEMS Microbiol. Lett. 204:75-79(2001).
[2]	
RN	NUCLEOTIDE SEQUENCE.
RP	STRAIN=RM1374;
RC	MEDLINE=22522600; PubMed=12635929;
RX	Rumer L., Jores J., Kirsch P., Cavaignac Y., Zehnke K., Wieler L.H.;
RA	"Dissemination of phev and phev located genomic islands among
RT	enteropathogenic (PEEC) and enterohemorrhagic (EHEC) E. coli and their
RT	possible role in the horizontal transfer of the locus of enterocyte
RT	effacement (LEE).";
RL	Int. J. Med. Microbiol. 292:463-475(2003).
[3]	
RN	NUCLEOTIDE SEQUENCE.
RP	STRAIN=RM1374;
RC	Jores J., Wagner S.K., Rumer L., Eichberg J., Laturnus C., Kirsch P.,
RA	Schierack P., Tschape H., Wieler L.H.;
RT	"Description of a 111-kb pathogenicity island (PAI) encoding various
RT	virulence features in the enterohemorrhagic E. coli (EHEC) strain
RT	RM1374 (O103:H2) and detection of a similar PAI in other EHEC strains
RT	of serotype O103:H2.";
RL	Int. J. Med. Microbiol. 294:417-425(2005).
CC	Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC	Distributed under the Creative Commons Attribution-NoDerivs license
CC	
DR	EMBL; AJ303141; CA134867.1; -; Genomic_DNA.
DR	SNR; OSKSP9; 261-325.
DR	GO; GO:0004872; F:receptor activity; IEA.
DR	GO; GO:0007155; P:cell adhesion; IEA.
DR	InterPro; IPR005336; Tir_rcpl.
DR	Pfam; PF07489; Tir_receptor_C; 1.
DR	Pfam; PF03549; Tir_receptor_M; 1.
DR	Pfam; PF07490; Tir_receptor_N; 1.
DR	PRINTS; PRO1370; TENSINTIMINR.
KW	Receptor.
SQ	SEQUENCE 538 AA; 55482 MW; 7F05D83C6207F117 CRC64;
QY	Query Match 72.7%; Score 120; DB 2; Length 538; Best Local Similarity 82.1%; Pred. No. 2.2e-06;

```

Matches 23; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
Oy 1 PIGNIGNNVNGNHLIPPAPIPSQTDGA 28
Db 2 PIGNIGHNPVNRALIPPAPIPSQTDGA 29

RESULT 22
O58187_ECOLI PRELIMINARY; PRT; 538 AA.
AC O58187;
DT 26-APR-2005, integrated into UniProtKB/TrEMBL.
DT 26-APR-2005, sequence version 1.
DT 07-FEB-2006, entry version 6.
DE Translocated intimin receptor.
GN Name=tir;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15870341; DOI=10.1128/AEM.71.5.2511-2519.2005;
RA Higgins J.A., Belt K.T., Karns J.S., Russell-Anelli J., Shelton D.R.;
RT "tir" and stx-Positive Escherichia coli in Stream Waters in a
RT Metropolitan Area.";
RL Appl. Environ. Microbiol. 71:2511-2519(2005).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Hohn C., Karns J.S., Higgins J.A.;
RL Submitted (FEB-2005) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
DR EMBL; AY944738; AAX4731.1; -; Genomic_DNA.
DR SRR; OS8187; 261-325.
DR GO; GO:0004872; P:receptor activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR003536; TIR_rcpt.
DR Pfam; PF07489; TIR_receptor_C; 1.
DR Pfam; PF03549; TIR_receptor_M; 1.
DR Pfam; PF07490; TIR_receptor_N; 1.
DR PRINTS; PR01370; TRNSINTIMNR.
KW Receptor.
SQ
SEQUENCE 538 AA; 55455 MW; D28F5200F04A1890 CRC64;

Query Match 67.3%; Score 111; DB 2; Length 538;
Best Local Similarity 78.6%; Pred. No. 36-05;
Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
Oy 1 PIGNIGNNVNGNHLIPPAPIPSQTDGA 28
Db 2 PIGNIGHNPVNRALIPPAPIPSQTDGA 29

RESULT 23
O34G79_RHOA PRELIMINARY; PRT; 1554 AA.
AC O34G79;
DT 06-DEC-2005, integrated into UniProtKB/TrEMBL.
DT 06-DEC-2005, sequence version 1.
DT 07-FEB-2006, entry version 3.
DE Hypothetical protein.
GN ORFNames=RPCDRAFT_4567;
OS Rhodopseudomonas palustris BisB18.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Rhodopseudomonas.
OX NCBI_TaxID=316056;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=BisB18;
RG US DOE Joint Genome Institute (JGI-RGF);

```

```

RA Copeland A., Lucas S., Lapidus A., Barry K., Deter J.C., Glavina T.,
RA Hammon N., Ierant S., Pitluck S., Richardson P.;
RT "Sequencing of the draft genome and assembly of Rhodopseudomonas
RT palustris BisB18.";
RT Submitted (OCT-2005) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=BisB18;
RG US DOE Joint Genome Institute (JGI-ORNL);
RA Larimer F., Land M.;
RT "Annotation of the draft genome assembly of Rhodopseudomonas palustris
RT BisB18.";
RT Submitted (NOV-2005) to the EMBL/GenBank/DBJ databases.
CC -! CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
DR EMBL; ALUR0100001; EAP13698.1; -; Genomic_DNA.
DR Hypothetical protein.
KW SEQUENCE 1554 AA; 148744 MW; 604FF187A3177D1E CRC64;
SQ
SEQUENCE 1554 AA; 148744 MW; 604FF187A3177D1E CRC64;

Query Match 40.3%; Score 66.5; DB 2; Length 1554;
Best Local Similarity 57.7%; Pred. No. 36;
Matches 15; Conservative 2; Mismatches 8; Indels 1; Gaps 1;
Oy 2 IGNIGNNVNGNHLIPPAPIPSQTD 26
Db 1250 VGSIGNLVGSNLAANTPPAPSPSPTD 1275

RESULT 24
O7XWS7_ORYSA
ID O7XWS7_ORYSA PRELIMINARY; PRT; 1510 AA.
AC O7XWS7;
DT 01-OCT-2003, integrated into UniProtKB/TrEMBL.
DT 01-MAR-2004, sequence version 2.
DT 07-FEB-2006, entry version 14.
DE OSUNBA0091C12.5 protein.
DE OSUNBA0091C12.5;
GN Name=OSUNBA0091C12.5;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; BEP clade;
OC Eriatoidaeae; Oryzaeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22337377; PubMed=12447439; DOI=10.1038/nature01183;
RA Peng Q., Zhang Y., Hao P., Wang S., Fu G., Huang Y., Li Y., Zhu J.,
RA Liu Y., Hu X., Jia P., Zhang Y., Zhao Q., Ying K., Yu S., Tang Y.,
RA Weng Q., Zhang L., Lu Y., Mu Y., Lu Y., Zhang L.S., Yu Z., Fan D.,
RA Liu X., Lu T., Li C., Wu Y., Sun T., Lei H., Li T., Hu H., Guan J.,
RA Wu M., Zhang R., Zhou B., Chen Z., Chen L., Jin Z., Wang R., Yin H.,
RA Cai Z., Ren S., Lv G., Gu W., Zhu G., Tu Y., Jia J., Zhang Y.,
RA Chen J., Kang H., Chen X., Shao C., Sun Y., Hu Q., Zhang X., Zhang W.,
RA Wang L., Ding C., Sheng H., Gu J., Chen S., Ni L., Zhu F., Chen W.,
RA Lan L., Lai Y., Cheng Z., Gu M., Jiang J., Li J., Hong G., Xue Y.,
RA Han B.;
RT "Sequence and analysis of rice chromosome 4.";
RL Nature 420:316-320(2002).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
DR EMBL; AL662989; CAD39927.2; -; Genomic_DNA.
DR Gramene; O7XWS7; -;
DR GO; GO:0003779; F:actin binding; IEA.
DR GO; GO:0030036; P:actin cytoskeleton organization and biogenesis; IEA.
DR GO; GO:0016043; P:cell organization and biogenesis; IEA.
DR InterPro; IPR003104; FH2_actin_bd.
DR Pfam; PF02181; FH2; 1.

```

```

DR SMART; SM00498; FH2; 1.
SQ SEQUENCE 1510 AA; 165213 MW; D372867E72F1E3FA CRC64;

Query Match
Best Local Similarity 39.1%; Score 64.5; DB 2; Length 1510;
Matches 14; Conservative 0; Mismatches 8; Indels 5; Gaps 1;

Qy 1 PIGNLGNVGNHILPPAPLPSEQTDG 27
Db 1126 PIGGL-----GQHAPPPPLPPIGIG 1147

RESULT 25
ID 08H3A7 ORYSA PRELIMINARY; PRT; 980 AA.
AC 08H3A7; 08GT20;
DT 01-MAR-2003, integrated into UniProtKB/TrEMBL.
DT 21-FEB-2006, entry version 16.
DE Putative exportin, tRNA (Nuclear export receptor for tRNAs).
GN Name:P0616D06.103; Synonyms:OJ1003_H02.130;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; BEP clade;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsumoto T., Katayose Y.;
RT "Oryza sativa nipponbare (GAS) genomic DNA, chromosome 7, PAC
clone: P0616D06."
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RA NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsumoto T., Yamamoto K.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs license
CC -----
DR EMBL; AP005198; BAC16491.1; -; Genomic DNA.
DR EMBL; AP003700; BAC79501.1; -; Genomic DNA.
DR Gramene; Q8H3A7; -.
DR GO; GO:0005488; F:binding; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR InterPro; IPR011989; ARM-like.
KM Receptor; Repeat.
SQ SEQUENCE 980 AA; 108381 MW; 376D95A28E9584F0 CRC64;

Query Match
Best Local Similarity 37.6%; Score 62; DB 2; Length 980;
Matches 12; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 14 LPPAPLPSPQTDGAAR 30
Db 133 LPPSPPLPGLPTDMPAR 149

RESULT 26
ID 04FX35 LEIMA PRELIMINARY; PRT; 1644 AA.
AC 04FX35;
DT 30-AUG-2005, integrated into UniProtKB/TrEMBL.
DT 30-AUG-2005, sequence version 1.
DT 07-FEB-2006, entry version 3.
DE Hypothetical protein.
GN ORFNames=LMJ1013;
OS Leishmania major strain Friedlin.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=347515;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA STRAIN=Friedlin;

```

```

RX PubMed=16020728; DOI=10.1126/science.1112680;
RA Irene A.C., Peacock C.S., Worthey E.A., Murphy L., Aggarwal G.,
RA Berriman M., Sisk E., Rajandream M.A., Adlem E., Aert R., Anupama A.,
RA Apostolou Z., Attipoe P., Baason N., Bauser C., Beck A., Beverley S.M.,
RA Blanchettin G., Borzym K., Bothe G., Brusch C.V., Collins M.,
RA Cadag E., Chariot L., Clayton C., Coulson R.M.R., Collins A.,
RA Cruz A.K., Davies R.M., De Gaudenzi J., Dobson D.E., Duesterhoef A.,
RA Fazellina G., Foster N., Frasch A.C., Fraser A.C., Fuchs M., Gabl C.,
RA Goble A., Goffeau A., Harris D., Hertz-Fowler C., Hilbert H., Horn D.,
RA Huang Y., Klages S., Knights A., Kude M., Larke N., Litvin L.,
RA Lord A., Louie T., Marra M., Masny D., Matthews K., Michael S.,
RA Mouttram J.C., Mueller-Auer S., Munden H., Nelson S., Norbertczak H.,
RA Oliver K., O'neil S., Pentony M., Pohl T.M., Price C., Purnelle B.,
RA Quail M.A., Rabinowitsch E., Reinhardt R., Rieger M., Rinta J.,
RA Robben J., Robertson L., Ruiz J.C., Rutter S., Saunders D.,
RA Schaefer M., Schein J., Schwartz D.C., Seeger K., Seyler A., Sharp S.,
RA Shin H., Sivam D., Squares R., Squares S., Totsato V., Vogt C.,
RA Volckaert G., Wambut R., Warren T., Wedler H., Woodward J., Zhou S.,
RA Zimmermann W., Smith D.F., Blackwell J.M., Stuart K.D., Barrell B.G.,
RA Wyler P.J.;
RT "The genome of the kinetoplastid parasite, Leishmania major."
RL Science 309:436-442 (2005).
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs license
CC -----
DR EMBL; CP000081; AAZ14308.1; -; Genomic DNA.
KW Hypothetical protein.
SQ SEQUENCE 1644 AA; 168098 MW; 7ED89FECDD1DA53DE CRC64;

Query Match
Best Local Similarity 37.6%; Score 62; DB 2; Length 1644;
Matches 10; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

Qy 5 LGNNVGNHILPPAPLPSEQTDGA 29
Db 1407 VSDSLHGEPPLPAAPVPRISADGAS 1431

RESULT 27
ID 08GJN5 DICDI PRELIMINARY; PRT; 279 AA.
AC 08GJN5;
DT 01-JUN-2003, integrated into UniProtKB/TrEMBL.
DT 01-JUN-2003, sequence version 1.
DT 07-FEB-2006, entry version 15.
DE Similar to Dictyostelium discoideum (Slime mold). STAC protein
DE (Hypothetical protein).
GN ORFNames=DDB0168607;
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA STRAIN=AX4.
RC MEDLINE=22092622; PubMed=12097910; DOI=10.1038/nature00847;
RA Gloeckner G., Eichinger L., Szafirski K., Pachebat J.A.,
RA Bankier A.T., Dear P.H., Lehmann R., Baumgart C., Parra G.,
RA Abril J.F., Guigo R., Kumpf K., Tunggal B., Cox E.C., Quail M.A.,
RA Platzer M., Rosenthal A., Noegel A.A.;
RT "Sequence and analysis of chromosome 2 of Dictyostelium discoideum."
RL Nature 418:79-85 (2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA STRAIN=AX4;
RC Baumgart C.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RA STRAIN=AX4;
RX PubMed=15875012; DOI=10.1038/nature03481;
RA Eichinger L., Pachebat J.A., Gloeckner G., Rajandream M.A.,
RA Sugang R., Berriman M., Song J., Olsen R., Szafirski K., Xu Q.,

```



```

DT 01-MAR-2003, sequence version 1.
DT 07-FEB-2006, entry version 13.
DE Survival of motor neuron.
GN Name=SMN;
OS Felis silvestris catus (Cat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Felidae;
OC Felinae; Felis.
OX NCBT_TaxID=9685;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA He Q., Lowrie C., Shelton G.D., Murphy W., Menotti-Raymond M.,
RA Fyfe J.C.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs license
CC -----
DR EMBL; AY094503; AAM18209.1; -; mRNA.
DR HSSP; Q16637; 1G5V.
DR SMR; QBHYB8; 86-141.
DR GO; GO:0005737; Cytoplasm; IEA.
DR GO; GO:0005634; Cytosol; IEA.
DR GO; GO:0003723; rRNA binding; IEA.
DR GO; GO:0006397; rRNA processing; IEA.
DR GO; GO:0000245; P:spliceosome assembly; IEA.
DR InterPro; IPR010304; SMN.
DR InterPro; IPR002999; Tudor.
DR Pfam; PF06003; SMN; 1.
DR SMART; SM00333; TUDOR; 1.
DR PROSITE; PSS0304; TUDOR; 1.
SQ SEQUENCE 290 AA; 31326 MW; E4D15F8447A96AC CRC64;

Query Match 36.7%; Score 60.5; DB 2; Length 290;
Best Local Similarity 39.4%; Pred. No. 31;
Matches 13; Conservative 3; Mismatches 12; Indels 5; Gaps 1;

OY 3 GNLGNVNG-----NHLIP-PAPPLPSQTGGAAR 30
Db 175 GNKSNVSKATSWNSFLPPPPGAGIGPAK 207

RESULT 31
O7TPJ8 RAT PRELIMINARY; PRT; 589 AA.
AC O7TPJ8;
DT 01-OCT-2003, integrated into UniProtKB/TrEMBL.
DT 01-OCT-2003, sequence version 1.
DT 07-FEB-2006, entry version 7.
DE AC2-143.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Rattus.
OX NCBT_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Xu C.S., Li W.Q., Li Y.C., Yang K.J., Yan H.M., Chang C.F., Zhao L.F.,
RA Ma H., Wang L., Wang S.F., Han H.P., Wang G.P., Chai L.Q., Yuan J.Y.,
RA Shi J.B., Rahman S., Wang Q.N., Zhang J.B.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs license
CC -----
DR EMBL; AY321337; AAP86269.1; -; mRNA.
DR Ensembl; ENSRNOG0000030738; Rattus norvegicus.
SQ SEQUENCE 589 AA; 65537 MW; CE569D7608F2C884 CRC64;

Query Match 36.7%; Score 60.5; DB 2; Length 589;
Best Local Similarity 40.5%; Pred. No. 68;
Matches 15; Conservative 5; Mismatches 10; Indels 7; Gaps 2;

```

```

OY 1 PIGNLGNVNG-----NHLIP-PAPPLPSQTGGAAR 30
Db 197 PRGNLGLGSGGTELSPOHSIPSPFPPTIAEGPR 233

RESULT 32
O5K9X9 CRYNE PRELIMINARY; PRT; 712 AA.
ID O5K9X9_CRYNE
AC O5K9X9;
DT 15-FEB-2005, integrated into UniProtKB/TrEMBL.
DT 15-FEB-2005, sequence version 1.
DT 07-FEB-2006, entry version 13.
DE Cytoplasm protein, putative.
GN OrderedLocustNames=CNK00330;
OS Cryptococcus neoformans (Filobasidiella neoformans).
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes;
OC Tremellomycetidae; Tremellales; Tremellaceae; Filobasidiella.
OX NCBT_TaxID=5207;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=JEC21;
RX PubMed=15653466; DOI=10.1126/science.1103773;
RA Loftus B.J., Fung E., Roncaglia P., Rowley D., Amedeo P., Bruno D.,
RA Vamathavan J., Miranda M., Anderson I.J., Fraser J.A., Allen J.E.,
RA Bosdet I.E., Brent M.R., Chiu R., Doering T.L., Donlin M.J.,
RA D'Souza C.A., Fox D.S., Grinberg V., Fu J., Fukushima M., Haas B.J.,
RA Huang J.C., Janson G., Jones S.J.M., Koo H.L., Krzywinski M.I.,
RA Kwon-Chung K.J., Lengler K.B., Maltz R., Maira M.A., Maira R.E.,
RA Mathewson C.A., Mitchell T.G., Petrea M., Riggs F.R., Salberg S.L.,
RA Schein J.E., Shvartsbeyn A., Shin H., Shumway M., Specht C.A.,
RA Suh B.B., Tenney A., Utecherback T.R., Wickes B.L., Wortman J.R.,
RA Wye N.H., Kronstad J.W., Lodge J.K., Heitman J., Davis R.W.,
RA Frazer C.M., Hyman R.W.;
RT "The genome of the basidiomycetous yeast and human pathogen
RT Cryptococcus neoformans."
RL Science 307:1321-1324 (2005).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs license
CC -----
DR EMBL; AE017351; AAM46085.1; -; Genomic DNA.
DR GO; GO:0003676; rNucleic acid binding; IEA.
DR GO; GO:0000166; rNucleotide binding; IEA.
DR InterPro; IPR012677; a_b_pla1c_nuc_bd.
DR InterPro; IPR001374; R3H_ss_bd.
DR InterPro; IPR001424; R3H_1.
DR Pfam; PF01424; R3H; 1.
DR Pfam; PF00076; RRM_1; 1.
DR SMART; SM00360; RRM; 1.
DR PROSITE; PSS1061; R3H; 1.
DR PROSITE; PSS0102; RRM; 1.
KW Complete proteome.
SQ SEQUENCE 712 AA; 75498 MW; 96D9A547542BDEC8 CRC64;

Query Match 36.7%; Score 60.5; DB 2; Length 712;
Best Local Similarity 61.9%; Pred. No. 85;
Matches 13; Conservative 1; Mismatches 4; Indels 3; Gaps 1;

OY 3 GNLGNVNGNHLIP-PAPPLPS 23
Db 503 GLFGNSVND---IPVPPLPS 520

RESULT 33
O5JUM3 CRYNE PRELIMINARY; PRT; 718 AA.
ID O5JUM3_CRYNE
AC O5JUM3;
DT 24-MAY-2005, integrated into UniProtKB/TrEMBL.
DT 24-MAY-2005, sequence version 1.
DT 07-FEB-2006, entry version 9.
DE Hypothetical protein.
GN ORFNames=CNBK3100;
OS Cryptococcus neoformans var. neoformans B-3501A.

```

```
CC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes;
OC Tremellomycetidae; Tremellales; Tremellaceae; Filobasidiella.
OX NCBI_TaxID=283643;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=B-3501A;
RA Fung E., Hyman R.W., Rowley D., Bruno D., Miranda M., Fukushima M.,
RA Wickes B.L., Fu J., Davis R.W.;
RT "Cryptococcus neoformans serotype D sequencing."
RL Submitted (JUL-2004) to the EMBL/Genbank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/Genbank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC distributed under the Creative Commons Attribution-NonDerivs License
CC -----
DR EMBL; AAEY01000053; EAU1959.1; -; Genomic DNA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:000166; F:nucleotide binding; IEA.
DR InterPro; IPR012677; a_b_plat_nuc_bd.
DR InterPro; IPR001374; R3H_ss_bd.
DR InterPro; IPR000504; RNPI_RNA_bd.
DR Pfam; PF01424; R3H_1.
DR Pfam; PF00076; RRM_1.
DR SMART; SM00360; RRM; 1.
DR PROSITE; PS51061; R3H; 1.
DR PROSITE; PS50102; RRM; 1.
KW Hypothetical protein.
SQ SEQUENCE 718 AA; 76198 MW; B259B93E04F4C0A3 CRC64;

Query Match 36.7%; Score 60.5; DB 2; Length 718;
Best Local Similarity 61.9%; Pred. No. 85;
Matches 13; Conservative 1; Mismatches 4; Indels 3; Gaps 1;

OY 3 GNLSNNVNGNHLIPAPLPSPS 23
Db 508 GLFGNSVND--IPVPPLPS 525

RESULT 34
OSB437_EMENT PRELIMINARY; PRT; 913 AA.
ID OSB437_EMENT
AC OSB437;
DT 26-APR-2005, integrated into UniProtKB/TrEMBL.
DT 26-APR-2005, sequence version 1.
DT 07-MAR-2006, entry version 5.
DE Hypothetical protein.
GN ORFNames=AN4693.2;
OS Aspergillus nidulans FGSC A4.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; Emmentella.
OX NCBI_TaxID=227321;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=FGSC 4;
RX PubMed=16372000; DOI=10.1038/nature04341;
RA Galagan J.E., Calvo S.E., Cuomo C., Ma L.-J., Wortman J.R.,
RA Batzoglou S., Lee S.-I., Basteniermen M., Spevak C.C., Clutterbuck J.,
RA Kucletsov V., Jutka J., Scanzocchio C., Farnham M., Butler J.,
RA Purcell S., Harris S., Braus G.H., Drah O., Busch S., D'Enfert C.,
RA Bouchier C., Goldman G.H., Bell-Pedersen D., Griffiths-Jones S.,
RA Doonan J.H., Yu J., Vienken K., Pain A., Freitag M., Selker E.U.,
RA Archer D.B., Penliva M.A., Oakley B.R., Momany M., Tanaka T.,
RA Kumagai T., Asai K., Machida M., Nierman W.C., Denning D.W.,
RA Kadacki M., Hynes W., Pooletti M., Fischer R., Miller B.L., Dyer P.S.,
RA Sachs M.S., Osmari S.A., Birren B.W.;
RT "Sequencing of Aspergillus nidulans and comparative analysis with A.
RT fumigatus and A. oryzae."
RT Nature 438:1105-1115(2005).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/Genbank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
```

```
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC distributed under the Creative Commons Attribution-NonDerivs License
CC -----
DR EMBL; AACD01000080; EAA60735.1; -; Genomic DNA.
KW Hypothetical protein.
SQ SEQUENCE 913 AA; 99021 MW; 8994A25FADB92FCD CRC64;

Query Match 36.7%; Score 60.5; DB 2; Length 913;
Best Local Similarity 40.5%; Pred. No. 1.1e+02;
Matches 17; Conservative 2; Mismatches 10; Indels 13; Gaps 3;

OY 1 PIGN-IGNNVNGNHLIPAPLPSPGTDGA 29
Db 689 PLPNPYIGNTNGPHDPGAGEQYPMPTYPAIPGQTPAPA 730

RESULT 35
OSYKX2_DROME PRELIMINARY; PRT; 613 AA.
ID OSYKX2_DROME
AC OSYKX2;
DT 01-MAY-2000, integrated into UniProtKB/TrEMBL.
DT 01-MAY-2000, sequence version 1.
DT 21-FEB-2006, entry version 22.
DE CG4636-PA (SD02991p) (SCAR).
GN Name=SCAR; Synonyms=Scar; ORFNames=CG4636, Dmel_CG4636;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miklos G.L.G.,
RA Abri1 J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Baer A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.B., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Folsler C., Gabriellian A.E., Gary N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laeko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson J.,
RA Mestl G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusken D.R., Pacib J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shie B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stappleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter B., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zhang X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RT Science 287:2185-2195(2000).
RN [2]
```

RP NUCLEOTIDE SEQUENCE.
 RA MEDLINE=22426065; PubMed=12537568;
 RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
 RA Patel S., Adams M., Champagne M., Dugan S.P., Frise E., Hodgson A.,
 RA George R.A., Hoskins R.A., Laverly T., Muzny D.M., Nelson C.R.,
 RA Pacled J.W., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
 RA Svitek R., Taber P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
 RA Weintraub G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
 RT "Finishing a whole-genome shotgun: release 3 of the Drosophila
 RT melanogaster euchromatic genome sequence";
 RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=22426070; PubMed=12537573;
 RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J.W., Svitek R.,
 RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
 RA Ashburner M., Celniker S.E.;
 RT "The transposable elements of the Drosophila melanogaster euchromatin:
 RT a genomic perspective";
 RL Genome Biol. 3:RESEARCH0084.1-RESEARCH0084.20(2002).
 RN [4]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=22426069; PubMed=12537572;
 RA Mista S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
 RA Hradecky P., Huang Y., Kaminker J.S., Milburn G.H., Prochnik S.E.,
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
 RA Betencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
 RA Lewis S.E.;
 RT "Annotation of the Drosophila melanogaster euchromatic genome: a
 RT systematic review";
 RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
 RN [5]
 RP NUCLEOTIDE SEQUENCE.
 RG Berkeley Drosophila Genome Project;
 RA Celniker S., Carlson J., Man K., Pfeiffer B., Frise E., George R.,
 RA Hoskins R., Stapleton M., Pacled J., Park S., Svitek R., Smith E.,
 RA Yu C., Rubin G.;
 RT "Drosophila melanogaster release 4 sequence";
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP NUCLEOTIDE SEQUENCE.
 RG FlyBase;
 RL Submitted (JAN-2006) to the EMBL/GenBank/DBJ databases.
 RN [7]
 RP NUCLEOTIDE SEQUENCE.
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champagne M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
 RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
 RA Nuno J., Pacled J., Parag V., Park S., Phouanavong S., Wan K.,
 RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 RN [8]
 RP NUCLEOTIDE SEQUENCE.
 RA Dickson B.J.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC distributed under the Creative Commons Attribution-NoDerivs license
 CC -----
 CC EMBL: AE003630; AAF53042.1; -; Genomic DNA.
 CC EMBL: AY061573; AAL29127.1; -; mRNA.
 CC EMBL: AF247763; AAF74194.1; -; mRNA.
 CC FlyBase: FBgn0041781; SCAR.
 DR GO:0030037; P-actin filament reorganization during cell c. .; IMP.
 DR GO:0007413; Paxonal fasciculation; IMP.
 DR GO:0030031; P-cell projection biogenesis; IMP.
 DR GO:0001745; P-compound eye morphogenesis (sensu Endoptery. .; IMP.
 DR GO:0030866; P-cortical actin cytoskeleton organization an. .; IMP.
 DR GO:0008335; P-ovarian ring canal stabilization; IMP.
 DR GO:0008360; P-regulation of cell shape; IMP.
 DR InterPro: IPR003124; WH2_actin_bd.

DR Pfam; PF02205; WH2; 1.
 DR SMART; SM00246; WH2; 1.
 DR PROSITE; PSS1082; WH2; 1.
 SQ SEQUENCE 613 AA; 67055 MW; C81578B9FC3F2B95 CRC64;
 Query Match 36.4%; Score 60; DB 2; Length 613;
 Best Local Similarity 47.8%; Pred. No. 83;
 Matches 11; Conservative 3; Mismatches 9; Indels 0; Gaps 0;
 QY 2 IGMIGNNVGNHLPAPPPPSQ 24
 DB 449 IGMIGNNVGNHLPAPPPPPVPPQ 471
 RESULT 36
 Q7YVA2_CRYPV PRELIMINARY; PRT; 938 AA.
 AC Q7YVA2;
 DT 01-OCT-2003, integrated into UniProtKB/TrEMBL.
 DT 01-OCT-2003, sequence version 1.
 DT 07-FEB-2006, entry version 11.
 DE Aob567, aof1001, aoe110, aoe264 and aoe130 genes, possible (large low
 DE complexity with PFL0360C pfal like Znfinger).
 GN ORFNames=1MB.355, cgd6.2950;
 OS Cryptosporidium parvum;
 OC Eukaryota, Alveolata, Apicomplexa, Coccidia, Eimeriida;
 OC Cryptosporidiidae; Cryptosporidium.
 OX NCBI_TaxId=5807;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=Iowa;
 RA Bankier A.T., Spriggs H.F., Fattmann B., Konfortov B.A., Madera M.,
 RA Vogel C., Teichmann S.A., Ivens A., Dear P.H.;
 RT "Integrated mapping, chromosomal sequencing and sequence analysis of
 RT Cryptosporidium parvum";
 RL Genome Res. 0:0-0(2003).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=Iowa type II;
 RX PubMed=15044751; DOI=10.1126/science.1094786;
 RA Abrahamson M.S., Templeton T.J., Enomoto S., Abraham J.E., Zhu G.,
 RA Lancio C.A., Deng M., Liu C., Widmer G., Tzipori S., Buck G.A., Xu P.,
 RA Bankier A.T., Dear P.H., Konfortov B.A., Spriggs H.F., Iyer L.,
 RA Anantharaman V., Aravind L., Kapur V.;
 RT "Complete genome sequence of the apicomplexan, Cryptosporidium
 RT parvum";
 RL Science 304:441-445(2004).
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC distributed under the Creative Commons Attribution-NoDerivs license
 CC -----
 CC EMBL: BX538351; CAD98583.1; -; Genomic DNA.
 CC EMBL: AAEE01000002; EAK89927.1; -; Genomic DNA.
 DR EMBL: AE00100002; EAK89927.1; -; Genomic DNA.
 SQ SEQUENCE 938 AA; 99600 MW; 2C103465B456D5C6 CRC64;
 Query Match 36.4%; Score 60; DB 2; Length 938;
 Best Local Similarity 62.5%; Pred. No. 1.3e+02;
 Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
 QY 13 HLPAPPPPSQTDGA 28
 DB 569 YLVPPPPPPSKTSGA 584
 RESULT 37
 Q726D4_DESVH
 ID Q726D4_DESVH PRELIMINARY; PRT; 974 AA.
 AC Q726D4;
 DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.
 DT 05-JUL-2004, sequence version 1.
 DT 07-FEB-2006, entry version 12.
 DE DNA internalization-related competence protein ComEC/Rec2.
 GN OrderedLocustNames=DVU3257; ORFNames=DVU_3257;


```

OS Desulfotribrio vulgaris (strain Hildenborough / ATCC 29579 / NCIMB
OS 8303).
OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfotribriales;
OC Desulfotribriaceae; Desulfotribrio.
CX NCBI_TaxID=882;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX PubMed=15077118; DOI=10.1038/nbt959;
RA Heidelberg J.F., Seebach R., Haveman S.A., Hemme C.L., Paulsen I.T.,
RA Kolonay J.F., Eisen J.A., Ward N.L., Methe B.A., Brinkac L.M.,
RA Daugherty S.C., DeBoy R.T., Dodson R.J., Durkin A.S., Madupu R.,
RA Nelson W.C., Sullivan S.A., Fouts D.E., Haft D.H., Selengut J.,
RA Peterson J.D., Davidson T.M., Zafar N., Zhou L., Radune D.,
RA Dimitrov G., Hance M., Tran K., Khouri H.M., Gill J., Uitterback T.R.,
RA Feldblyum T.V., Wall J.D., Voordouw G., Fraser C.M.;
RT "The genome sequence of the anaerobic, sulfate-reducing bacterium
RT Desulfotribrio vulgaris Hildenborough.";
RL Nat. Biotechnol. 22:554-559(2004).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs license
CC -----
DR EMBL: AE017285; AAS97727.1; -; Genomic_DNA.
DR TIGR: DVU3257; -.
DR BioCyc: DVUL882:DVU3257-MONOMER; -.
DR InterPro: IPR001279; Blactamase-like.
DR InterPro: IPR004477; ComEC_N-term.
DR Pfam: PF03772; Competence_1.
DR Pfam: PF00753; Lactamase_B_1.
DR TIGRfams: TIGR00360; ComEC_N-term; 1.
DR Complete proteome.
SQ SEQUENCE 974 AA; 102101 MW; 898B5D40904E2D39 CRC64;

Query Match 36.4%; Score 60; DB 2; Length 974;
Best Local Similarity 50.0%; Pred. No. 14e+02;
Matches 11; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

QY 3 GNLGNVNGNHLIPAPLPSPQ 24
Db 189 GGTGCGVAPALVPAPLPAPLPGR 210

RESULT 38
O2WIP9 MAGSA PRELIMINARY; PRT; 9529 AA.
AC O2WIP9;
DT 10-JAN-2006, integrated into UniProtKB/TrEMBL.
DT 10-JAN-2006, sequence version 1.
DT 07-FEB-2006, entry version 3.
DE Hypothetical protein.
GN ORFNames=amb3422;
OS Magnetospirillum magneticum AMB-1.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodospirillales;
OC Rhodospirillaceae; Magnetospirillum.
OX NCBI_TaxID=342108;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX STRAIN=AMB-1;
RX PubMed=16303747;
RA Matsunaga T., Okamura Y., Fukuda Y., Wakiyudi A.T., Murase Y.,
RA Takeyama H.;
RT "Complete Genome Sequence of the Facultative Anaerobic Magnetotactic
RT Bacterium Magnetospirillum sp. strain AMB-1.";
RL DNA Res. 12:157-166(2005).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs license
CC -----
DR EMBL: AP007255; BAE5226.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 9529 AA; 955916 MW; 69B375E228915BD4 CRC64;

Query Match 36.4%; Score 60; DB 2; Length 9529;

```

```

Best Local Similarity 48.3%; Pred. No. 1.8e+03;
Matches 14; Conservative 3; Mismatches 10; Indels 2; Gaps 1;

QY 1 PIGNLGNVNGNHLIPAPLPSPQTDGA 29
Db 6108 PVSSL-NANGPSLPIGNMSPSPDQGA 6134

RESULT 39
O2QM44 ORYSA PRELIMINARY; PRT; 329 AA.
AC O2QM44;
DT 24-JAN-2006, integrated into UniProtKB/TrEMBL.
DT 24-JAN-2006, sequence version 1.
DT 07-MAR-2006, entry version 3.
DE ZF-HD protein dimerization region, putative.
GN ORFNames=LOC_0812g10630;
OS Oryza sativa (Japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; BEP clade;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=33947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Buehl C.R., Wing R.A., McCombie W.A., Ouyang S.;
RL Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs license
CC -----
DR EMBL: DP000011; ABA96146.1; -; Genomic_DNA.
KW DNA-binding.
SQ SEQUENCE 329 AA; 34952 MW; B4822362A5CB63D6 CRC64;

Query Match 35.8%; Score 59; DB 2; Length 329;
Best Local Similarity 38.5%; Pred. No. 55;
Matches 10; Conservative 4; Mismatches 12; Indels 0; Gaps 0;

QY 3 GNLGNVNGNHLIPAPLPSPQTDGA 28
Db 35 GGGGGGNGNRHMSPTPPAALBESKA 60

RESULT 40
O4D3Y5 TRYCR PRELIMINARY; PRT; 303 AA.
AC O4D3Y5;
DT 13-SEP-2005, integrated into UniProtKB/TrEMBL.
DT 13-SEP-2005, sequence version 1.
DT 07-FEB-2006, entry version 2.
DE Hypothetical protein.
GN ORFNames=Tc00.1047053511461.50;
OS Trypanosoma cruzi.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma;
OC Schizotrypanum.
OX NCBI_TaxID=5693;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX STRAIN=CL Brenner;
RA El-Sayed N.M.A., Myler P.J., Bartholomew D.C., Nilsson D.,
RA Aggarwal G., Tran A.-N., Ghedin E., Worthey E.A., Delcher A.L.,
RA Blandin G., Westenberg S.J., Caler E., Cerqueira G.C., Branche C.,
RA Haas B., Anapuma A., Arner E., Aislund L., Attipoe P., Bontempo E.,
RA Bringaud F., Burton P., Cadag E., Campbell D.A., Carlington M.,
RA Crastire J., Darban H., da Silveira U.F., de Jong P., Edwards K.,
RA Englund P.T., Fazelina G., Feldblyum T., Ferella M., Frauch A.C.,
RA Gull K., Horn D., Hou L., Huang Y., Kindlund E., Klingbeil M.,
RA Kluge S., Koo H., Lacerda D., Levin M.J., Lorenzi H., Louie T.,
RA Machado C.R., McCulloch R., McKenna A., Mizuno Y., Motttram J.C.,
RA Nelson S., Ochaya S., Osoegawa K., Pai G., Parsons M., Pentony M.,
RA Peterson U., Pop M., Ramirez J.L., Rinta J., Robertson L.,
RA Salzberg S.L., Sanchez D.O., Seyler A., Sharma R., Shetty J.,
RA Simpson A.J., Slek B., Tammi M.T., Tarleton R., Teixeira S.,

```


RA Van Aken S., Voigt C., Ward P.N., Wickstead B., Wortman J., White O.,
 RA Fraser C.M., Stuart K.D., Andersson B.,
 RT "The Genome Sequence of *Trypanosoma cruzi*, Etiologic Agent of Chagas'
 RT Disease.",
 RL Science 0:0-0(2005) .
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=CL Brenner;
 RA El-Sayed N.M.A., Myler P.J., Blandin G., Berriman M., Crabtree J.,
 RA Aggarwal G., Caler E., Renauld H., Worthey E.A., Hertz-Fowler C.,
 RA Ghedin E., Peacock C., Bartholomew D.C., Haas B.J., Tran A.-N.,
 RA Wortman J.R., Altmark U.C.M., Angiuoli S., Anupama A., Badger J.,
 RA Brinkauf F., Cadog E., Carlton J.M., Cerqueira G.C., Creasy T.,
 RA Delcher A.L., Djikeng A., Embley T.M., Hauser C., Ivens A.C.,
 RA Kummerfeld S.K., Pereira-Leal J.B., Nilsson D., Peterson J.,
 RA Salzberg S.L., Shallom J., Silva J.C., Sundaram J., Westenberg S.,
 RA White O., Melville S.E., Donelson J.E., Andersson B., Stuart K.D.,
 RA Hall N.;
 RT "Comparative Genomics of *Trypanosomatid* Parasitic Protozoa.",
 RL Science 0:0-0(2005) .
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=CL Brenner;
 RA El-Sayed N., Bartholomew D., Haas B.;
 RL Submitted (JUN-2005) to the EMBL/Genbank/DBJ databases.
 CC -!- CAUTION: The sequence shown here is derived from an
 CC EMBL/Genbank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 CC -----
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NoDerivs license
 CC -----
 DR EMBL; AAHK01001066; EAA87241.1; -; Genomic_DNA.
 KW Hypothetical protein.
 SQ SEQUENCE 303 AA; 32793 MW; 93B64E9984B47259 CRC64;

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: August 1, 2006, 21:47:00 ; Search time 51 Seconds
(without alignments)
51.489 Million cell updates/sec

Title: US-09-189-415D-7
Perfect score: 165
Sequence: 1 PIGNLGNNVGNHILPPAPPLPQTDGAAR 30

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 650591 seqs, 87530628 residues

Total number of hits satisfying chosen parameters: 650591

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Maximum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : Issued Patents AA:*
1: /EMC_Celerra_SIDS3/ptodata/2/1aa/5_COMB.pep:*
2: /EMC_Celerra_SIDS3/ptodata/2/1aa/6_COMB.pep:*
3: /EMC_Celerra_SIDS3/ptodata/2/1aa/7_COMB.pep:*
4: /EMC_Celerra_SIDS3/ptodata/2/1aa/H_COMB.pep:*
5: /EMC_Celerra_SIDS3/ptodata/2/1aa/PCYTUS_COMB.pep:*
6: /EMC_Celerra_SIDS3/ptodata/2/1aa/RE_COMB.pep:*
7: /EMC_Celerra_SIDS3/ptodata/2/1aa/backfilltest.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	57	34.5	268	2	US-09-270-767-41663
2	56.5	34.2	22	2	US-08-630-915A-67
3	56.5	34.2	22	2	US-09-879-957-67
4	56	33.9	145	2	US-09-640-211A-794
5	56	33.9	145	2	US-09-489-039A-13282
6	55.5	33.6	674	2	US-08-653-648A-14
7	55.5	33.6	675	2	US-09-564-418-12
8	55	33.3	631	2	US-09-252-991A-26444
9	54.5	33.0	618	2	US-09-252-991A-26444
10	54	32.7	211	2	US-09-248-796A-22421
11	54	32.7	623	2	US-09-452-991A-30386
12	54	32.1	2736	2	US-09-252-991A-30227
13	53	32.1	1251	2	US-09-270-767-60141
14	53	32.1	253	2	US-09-270-767-44686
15	53	32.1	569	2	US-08-311-731A-106
16	53	32.1	832	2	US-09-252-991A-19252
17	53	31.5	219	2	US-09-452-991A-28228
18	52	31.5	366	2	US-09-377-285B-20
19	52	31.5	366	2	US-10-192-381-20
20	52	31.5	476	2	US-09-189-035-1
21	52	31.5	476	2	US-09-382-086-1
22	52	31.5	690	2	US-09-252-991A-23187
23	51	30.9	55	2	US-09-970-767-61121
24	51	30.9	509	2	US-10-104-047-2072
25	51	30.9	644	2	US-09-198-452A-63
26	51	30.9	1255	1	US-09-080-897-4

27	51	30.9	1255	2	US-08-899-595-1	Sequence 1, Appli
28	51	30.9	1255	2	US-09-323-735-4	Sequence 4, Appli
29	50.5	30.6	78	2	US-09-270-767-60585	Sequence 60585, A
30	50.5	30.6	212	2	US-09-692-945-2	Sequence 2, Appli
31	50.5	30.6	234	2	US-09-895-674A-1	Sequence 1, Appli
32	50.5	30.6	259	2	US-09-270-767-45089	Sequence 45089, A
33	50.5	30.6	347	2	US-09-623-497-1	Sequence 1, Appli
34	50.5	30.6	744	2	US-10-332-995-12	Sequence 12, Appli
35	50.5	30.6	1017	2	US-08-468-996-10	Sequence 10, Appli
36	50.5	30.6	1060	2	US-08-311-820-3	Sequence 3, Appli
37	50.5	30.6	1418	2	US-08-963-825-20	Sequence 20, Appli
38	50.5	30.6	1418	2	US-09-010-999-1	Sequence 1, Appli
39	50.5	30.6	1418	2	US-09-500-811-20	Sequence 20, Appli
40	50.5	30.6	1418	2	US-09-570-573-20	Sequence 20, Appli
41	50.5	30.6	1418	2	US-09-548-608-20	Sequence 20, Appli
42	50	30.3	128	2	US-10-104-047-3112	Sequence 3112, Ap
43	50	30.3	133	2	US-09-248-796A-23164	Sequence 23164, A
44	50	30.3	196	2	US-10-101-464A-981	Sequence 981, App
45	50	30.3	451	2	US-09-949-016-9567	Sequence 9567, Ap
46	50	30.3	496	2	US-09-949-016-6987	Sequence 6987, Ap
47	50	30.3	616	2	US-09-534-228B-1	Sequence 1, Appli
48	50	30.3	694	2	US-09-949-016-7117	Sequence 7117, Ap
49	50	30.3	1248	1	US-09-080-897-2	Sequence 2, Appli
50	50	30.3	1248	2	US-09-323-735-2	Sequence 2, Appli
51	50	30.3	1315	2	US-08-899-595-3	Sequence 3, Appli
52	49.5	30.0	51	2	US-08-545-196B-62	Sequence 62, Appli
53	49.5	30.0	51	3	US-09-109-082A-62	Sequence 62, Appli
54	49.5	30.0	90	2	US-09-489-039A-10146	Sequence 10146, A
55	49.5	30.0	170	5	PCT-US95-16806A-6	Sequence 6, Appli
56	49.5	30.0	278	2	US-08-545-196B-19	Sequence 19, Appli
57	49.5	30.0	278	3	US-09-109-082A-19	Sequence 19, Appli
58	49.5	30.0	294	1	US-08-612-986-6	Sequence 6, Appli
59	49.5	30.0	294	1	US-08-361-806A-6	Sequence 6, Appli
60	49.5	30.0	294	2	US-08-545-196B-9	Sequence 9, Appli
61	49.5	30.0	294	2	US-09-028-327-4	Sequence 4, Appli
62	49.5	30.0	294	2	US-09-571-078A-4	Sequence 4, Appli
63	49.5	30.0	294	2	US-09-109-082A-9	Sequence 9, Appli
64	49.5	30.0	294	5	PCT-US95-16806A-21	Sequence 21, Appli
65	49.5	30.0	304	2	US-09-949-016-8565	Sequence 8565, Ap
66	49.5	30.0	304	2	US-09-949-016-8566	Sequence 8566, Ap
67	49.5	30.0	304	2	US-09-949-016-8567	Sequence 8567, Ap
68	49.5	30.0	304	2	US-09-949-016-8568	Sequence 8568, Ap
69	49.5	30.0	304	2	US-09-949-016-8569	Sequence 8569, Ap
70	49.5	30.0	304	2	US-09-949-016-8570	Sequence 8570, Ap
71	49.5	30.0	316	2	US-09-949-016-7099	Sequence 7099, Ap
72	49.5	30.0	316	2	US-09-949-016-7100	Sequence 7100, Ap
73	49.5	30.0	316	2	US-09-949-016-7101	Sequence 7101, Ap
74	49.5	30.0	316	2	US-09-949-016-7102	Sequence 7102, Ap
75	49.5	30.0	316	2	US-09-949-016-7103	Sequence 7103, Ap
76	49.5	30.0	316	2	US-09-949-016-7104	Sequence 7104, Ap
77	49.5	30.0	319	2	US-09-252-991A-26784	Sequence 26784, A
78	49	29.7	78	2	US-09-248-796A-24951	Sequence 24951, A
79	49	29.7	130	2	US-10-101-464A-780	Sequence 780, App
80	49	29.7	254	2	US-09-449-437A-4	Sequence 4, Appli
81	49	29.7	408	2	US-09-949-016-9941	Sequence 9941, Ap
82	49	29.7	409	2	US-09-248-796A-21624	Sequence 21624, A
83	49	29.7	640	2	US-09-949-016-8010	Sequence 8010, Ap
84	49	29.7	699	2	US-10-104-047-2814	Sequence 2814, Ap
85	49	29.7	1088	2	US-09-949-016-6644	Sequence 6644, Ap
86	49	29.7	1147	2	US-09-949-016-8616	Sequence 8616, Ap
87	48.5	29.4	142	1	US-08-681-812-2	Sequence 2, Appli
88	48.5	29.4	211	2	US-09-692-945-4	Sequence 4, Appli
89	48.5	29.4	211	2	US-09-665-493B-5	Sequence 5, Appli
90	48.5	29.4	211	2	US-10-011-364-2	Sequence 2, Appli
91	48.5	29.4	242	2	US-09-602-777A-92	Sequence 92, Appli
92	48.5	29.4	249	2	US-09-949-016-8805	Sequence 8805, Ap
93	48.5	29.4	267	2	US-09-602-777A-90	Sequence 90, Appli
94	48.5	29.4	347	2	US-09-902-594A-15136	Sequence 15136, A
95	48.5	29.4	357	2	US-09-252-991A-26979	Sequence 26979, A
96	48.5	29.4	416	2	US-09-949-016-8977	Sequence 8977, Ap
97	48.5	29.4	516	2	US-09-949-002-556	Sequence 526, App
98	48.5	29.4	557	1	US-08-424-788-6	Sequence 63, Appli
99	48.5	29.4	578	1	US-08-424-788-5	Sequence 5, Appli

100 48.5 29.4 578 1 US-08-110-683-2 Sequence 2, Appli

ALIGNMENTS

RESULT 1

US-09-270-767-41663
; Sequence 41663, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 41663
; LENGTH: 268
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-41663

Query Match 34.5%; Score 57; DB 2; Length 268;

Best Local Similarity 48.1%; Pred. No. 19;
Matches 13; Conservative 2; Mismatches 10; Indels 2; Gaps 1;

Qy 3 GNLGNVNGNHL-IPPAPLPSQTDG 27

Db 155 GKNGFGASGDHLPQAPAPAPQVPTAG 181

RESULT 2

US-08-630-915A-67
; Sequence 67, Application US/08630915A
; Patent No. 6309820
; GENERAL INFORMATION:
; APPLICANT: SPARKS, Andrew B.
; APPLICANT: HOFFMAN, No. 6309820h
; APPLICANT: KAY, Brian K.
; APPLICANT: FOWLES, Dana M.
; APPLICANT: MCCONNELL, Stephen J.
; TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL
; TITLE OF INVENTION: DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND
; NUMBER OF SEQUENCES: 227
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/630,915A
; FILING DATE: 03-APR-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-174
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 PENNIE

; INFORMATION FOR SEQ ID NO: 67:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 22 amino acids

; TYPE: amino acid

; STRANDEDNESS:

; TOPOLOGY: unknown

; MOLECULE TYPE: peptide

; FEATURE:

; NAME/KEY: Other

; OTHER INFORMATION: Biotinylated N-terminal

US-08-630-915A-67

Query Match 34.2%; Score 56.5; DB 2; Length 22;

Best Local Similarity 50.0%; Pred. No. 1.4;

Matches 11; Conservative 3; Mismatches 5; Indels 3; Gaps 1;

Qy 6 GNNVNGNHLIPPAPLPSQTDG 27

Db 4 GNYVNA--LPPGPPLPKXNG 22

RESULT 3

US-09-879-957-67
; Sequence 67, Application US/09879957
; Patent No. 6709821
; GENERAL INFORMATION:
; APPLICANT: SPARKS, Andrew B.
; APPLICANT: HOFFMAN, No. 6709821h
; APPLICANT: KAY, Brian K.
; APPLICANT: FOWLES, Dana M.
; APPLICANT: MCCONNELL, Stephen J.
; TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL
; TITLE OF INVENTION: DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND
; USING SAME
; NUMBER OF SEQUENCES: 227
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/879,957
; FILING DATE: 13-Jun-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/630,915
; FILING DATE: 03-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-174
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 67:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Other
; OTHER INFORMATION: Biotinylated N-terminal
; SEQUENCE DESCRIPTION: SEQ ID NO: 67:

US-09-879-957-67

Query Match

34.2%; Score 56.5; DB 2; Length 22;
Best Local Similarity 50.0%; Pred. No. 1.4;
Matches 11; Conservative 3; Mismatches 5; Indels 3; Gaps 1;

QY

6 GNNVNGNHLPPAPPLPSQTDG 27

Db

4 GNYVNA---LPPGPPLPAKNG 22

RESULT 4

US-09-640-211A-794
; Sequence 794, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Modifications and Methods for the
; FILE REFERENCE: 11000.1021C1U
; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 794
; LENGTH: 145
; TYPE: PRT
; ORGANISM: Eucalyptus grandis
US-09-640-211A-794

Query Match

33.9%; Score 56; DB 2; Length 145;
Best Local Similarity 38.5%; Pred. No. 13;
Matches 10; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

QY

3 GNLGNVNGNHLPPAPPLPSQTDGA 28

Db

61 GGVDNNGGYLHSPVMPKSDGS 86

RESULT 5

US-09-489-039A-13282
; Sequence 13282, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 13282
; LENGTH: 474
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-13282

Query Match

33.9%; Score 56; DB 2; Length 474;
Best Local Similarity 55.0%; Pred. No. 47;
Matches 11; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY

11 GNNLPPAPPLPSQTDGAAR 30

Db

81 GNNLPPAPGLPMKQAPAR 100

RESULT 6

US-08-653-648A-14

; Sequence 14, Application US/08653648A

; Patent No. 6379945
; GENERAL INFORMATION:
; APPLICANT: Jepson, Ian
; APPLICANT: Greenland, Andrew
; APPLICANT: Martinez, Alberto
; TITLE OF INVENTION: A Gene Switch
; FILE REFERENCE: PPD50047/US
; CURRENT APPLICATION NUMBER: US/08/653,648A
; CURRENT FILING DATE: 1996-05-24
; PRIOR APPLICATION NUMBER: GB 9510759.5
; PRIOR FILING DATE: 1995-05-26
; PRIOR APPLICATION NUMBER: GB 9605656.9
; PRIOR FILING DATE: 1996-03-18
; PRIOR APPLICATION NUMBER: GB 9513882.2
; PRIOR FILING DATE: 1995-07-07
; PRIOR APPLICATION NUMBER: GB 9517316.7
; PRIOR FILING DATE: 1995-08-24
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 14
; LENGTH: 674
; TYPE: PRT
; ORGANISM: Aedes aegypti
US-08-653-648A-14

Query Match

33.6%; Score 55.5; DB 2; Length 674;
Best Local Similarity 52.0%; Pred. No. 80;
Matches 13; Conservative 2; Mismatches 9; Indels 1; Gaps 1;

QY

2 IGNLGNVNGNHLPPAPPLPSQT 25

Db

115 VGNLINGVNPOTLIPPLPSITQNT 139

RESULT 7

US-09-564-418-12
; Sequence 12, Application US/09564418
; Patent No. 6610828
; GENERAL INFORMATION:
; APPLICANT: Syngenta
; APPLICANT: Jepson, Ian
; APPLICANT: Martinez, Alberto
; APPLICANT: Greenland, Andrew James
; TITLE OF INVENTION: A GENE SWITCH
; FILE REFERENCE: 1392/4/3
; CURRENT APPLICATION NUMBER: US/09/564,418
; CURRENT FILING DATE: 2000-05-03
; PRIOR APPLICATION NUMBER: US 09/564,418
; PRIOR FILING DATE: 2000-05-03
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 12
; LENGTH: 675
; TYPE: PRT
; ORGANISM: Aedes aegypti
US-09-564-418-12

Query Match

33.6%; Score 55.5; DB 2; Length 675;
Best Local Similarity 52.0%; Pred. No. 80;
Matches 13; Conservative 2; Mismatches 9; Indels 1; Gaps 1;

QY

2 IGNLGNVNGNHLPPAPPLPSQT 25

Db

115 VGNLINGVNPOTLIPPLPSITQNT 139

RESULT 8

US-09-252-991A-26444
; Sequence 26444, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rudenfield et al.

Db 559 PIGOTTADANGNWSFTPTPLPDGT 583

RESULT 13

US-09-270-767-60141
Sequence 60141, Application US/09270767
Patent No. 6703491
GENERAL INFORMATION:
APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270,767
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 60141
LENGTH: 121
TYPE: PRT
ORGANISM: Drosophila melanogaster
US-09-270-767-60141

Query Match 32.1%; Score 53; DB 2; Length 121;
Best Local Similarity 34.4%; Pred. No. 24;
Matches 11; Conservative 7; Mismatches 10; Indels 4; Gaps 1;

QY 1 PIGNGNVNGNHLPPAPLP-----SQTGDA 28

Db 43 PKSSRRSRISGGLVMPMPVPMIVGQGDGA 74

RESULT 14

US-09-270-767-44686
Sequence 44686, Application US/09270767
Patent No. 6703491
GENERAL INFORMATION:
APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270,767
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 44686
LENGTH: 253
TYPE: PRT
ORGANISM: Drosophila melanogaster
FEATURE:
OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-44686

Query Match 32.1%; Score 53; DB 2; Length 253;
Best Local Similarity 34.4%; Pred. No. 55;
Matches 11; Conservative 7; Mismatches 10; Indels 4; Gaps 1;

QY 1 PIGNGNVNGNHLPPAPLP-----SQTGDA 28

Db 175 PKSSRRSRISGGLVMPMPVPMIVGQGDGA 206

RESULT 15

US-08-311-731A-106
Sequence 106, Application US/08311731A
Patent No. 658326
GENERAL INFORMATION:
APPLICANT: SMITH, DOUGLAS
APPLICANT: MAO, JEN-I
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
TITLE OF INVENTION: RELATING TO MYCOBACTERIUM TUBERCULOSIS AND LAPRAE FOR
NUMBER OF SEQUENCES: 411
CORRESPONDENCE ADDRESS:
ADDRESSEE: WOLF, GREENFIELD & SACKS, P.C.
STREET: 600 ATLANTIC AVENUE

CITY: BOSTON
STATE: MASSACHUSETTS
COUNTRY: USA
ZIP: 02210

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/311,731A
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: GATES, EDWARD R.
REGISTRATION NUMBER: 31,616
REFERENCE/DOCKET NUMBER: C0044/7125
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/720-3500
TELEFAX: 617/720-2441

INFORMATION FOR SEQ ID NO: 106:

SEQUENCE CHARACTERISTICS:

LENGTH: 569 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: YES

ORIGINAL SOURCE:

ORGANISM: MYCOBACTERIUM LEPRAE

US-08-311-731A-106

Query Match 32.1%; Score 53; DB 2; Length 569;
Best Local Similarity 64.3%; Pred. No. 1,4e+02;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 16 PPAPLPQSQTGDA 29

Db 183 PPAPLPQHTDGT 196

RESULT 16

US-09-252-991A-19252
Sequence 19252, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 19252
LENGTH: 832
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19252

Query Match 32.1%; Score 53; DB 2; Length 832;
Best Local Similarity 45.5%; Pred. No. 2.1e+02;
Matches 10; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

QY 1 PIGNGNVNGNHLPPAPLP 22

Db 406 PALGRHRRAGRHLPAPRRP 427

RESULT 17

US-09-252-991A-28228

```
; Sequence 28228, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 28228
; LENGTH: 219
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28228

Query Match          31.5%; Score 52; DB 2; Length 219;
Best Local Similarity 50.0%; Pred. No. 63;
Matches 9; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

OY 13 PPAPPLPSQTDGAR 30
Db 11 HVRPPCPPLPPQRPASPR 28

RESULT 18
US-09-377-285B-20
; Sequence 20, Application US/09377285B
; Patent No. 6720175
; GENERAL INFORMATION:
; APPLICANT: THE JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE
; APPLICANT: WORLEY, Paul
; APPLICANT: TU, Jian
; APPLICANT: XIAO, Bo
; APPLICANT: LEAHY, Daniel
; APPLICANT: BENKEN, Jutta
; APPLICANT: LANAHAN, Anthony
; TITLE OF INVENTION: NUCLEIC ACID MOLECULE ENCODING HOMER 1b PROTEIN (AS AMENDED)
; FILE REFERENCE: JHU1580-4
; CURRENT APPLICATION NUMBER: US/09/377,285B
; CURRENT FILING DATE: 1999-08-18
; PRIOR APPLICATION NUMBER: US 60/138,426
; PRIOR FILING DATE: 1999-06-10
; PRIOR APPLICATION NUMBER: US 60/138,493
; PRIOR FILING DATE: 1999-06-10
; PRIOR APPLICATION NUMBER: US 60/138,494
; PRIOR FILING DATE: 1999-06-10
; PRIOR APPLICATION NUMBER: US 60/097,334
; PRIOR FILING DATE: 1998-08-18
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 20
; LENGTH: 366
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-377-285B-20

Query Match          31.5%; Score 52; DB 2; Length 366;
Best Local Similarity 81.8%; Pred. No. 1.1e+02;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 16 PPAPPLPSQTD 26
Db 232 PPAPPLPSSLD 242

RESULT 19
US-10-192-381-20
; Sequence 20, Application US/10192381
; Patent No. 6864083
; GENERAL INFORMATION:
; APPLICANT: THE JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE
; APPLICANT: WORLEY, Paul
; APPLICANT: TU, Jian
; APPLICANT: XIAO, Bo
; APPLICANT: LEAHY, Daniel
; APPLICANT: BENKEN, Jutta
; APPLICANT: LANAHAN, Anthony
; TITLE OF INVENTION: NUCLEIC ACID MOLECULE ENCODING HOMER 1b PROTEIN (AS AMENDED)
; FILE REFERENCE: JHU1580-4
; CURRENT APPLICATION NUMBER: US/10/192,381
; CURRENT FILING DATE: 2002-07-09
; PRIOR APPLICATION NUMBER: US/09/377,285
; PRIOR FILING DATE: 1999-08-18
; PRIOR APPLICATION NUMBER: US 60/138,426
; PRIOR FILING DATE: 1999-06-10
; PRIOR APPLICATION NUMBER: US 60/138,493
; PRIOR FILING DATE: 1999-06-10
; PRIOR APPLICATION NUMBER: US 60/138,494
; PRIOR FILING DATE: 1999-06-10
; PRIOR APPLICATION NUMBER: US 60/097,334
; PRIOR FILING DATE: 1998-08-18
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 20
; LENGTH: 366
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-192-381-20

Query Match          31.5%; Score 52; DB 2; Length 366;
Best Local Similarity 81.8%; Pred. No. 1.1e+02;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 16 PPAPPLPSQTD 26
Db 232 PPAPPLPSSLD 242

RESULT 20
US-09-189-035-1
; Sequence 1, Application US/09189035
; Patent No. 6020165
; GENERAL INFORMATION:
; APPLICANT: Yue, Henry
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl J.
; APPLICANT: Baughn, Mariah R.
; TITLE OF INVENTION: CYTOKINE SIGNAL REGULATORS
; FILE REFERENCE: PF-0638 US
; CURRENT APPLICATION NUMBER: US/09/189,035
; CURRENT FILING DATE: 1998-11-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PERL Program
; SEQ ID NO 1
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE: -
; OTHER INFORMATION: 2280326
US-09-189-035-1

Query Match          31.5%; Score 52; DB 2; Length 476;
Best Local Similarity 81.8%; Pred. No. 1.5e+02;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 16 PPAPPLPSQTD 26
Db 342 PPAPPLPSSLD 352
```



```
RESULT 21
US-09-382-086-1
; Sequence 1, Application US/09382086
; Patent No. 6201106
; GENERAL INFORMATION:
; APPLICANT: Yue, Henry
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl J.
; APPLICANT: Baughn, Mariah R.
; TITLE OF INVENTION: CYTOKINE SIGNAL REGULATORS
; FILE REFERENCE: PF-0638 US
; CURRENT APPLICATION NUMBER: US/09/382,086
; CURRENT FILING DATE: 1999-08-24
; EARLIER APPLICATION NUMBER: 09/189,035
; EARLIER FILING DATE: 1998-11-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PERL Program
; SEQ ID NO 1
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE: -
; OTHER INFORMATION: 2280326
US-09-382-086-1

Query Match          31.5%; Score 52; DB 2; Length 476;
Best Local Similarity 81.8%; Pred. No. 1.5e+02;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      16 PPAPPLPSQTD 26
Db      342 PPAPPLPSSLD 352

RESULT 22
US-09-252-991A-23187
; Sequence 23187, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 23187
; LENGTH: 690
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23187

Query Match          31.5%; Score 52; DB 2; Length 690;
Best Local Similarity 33.3%; Pred. No. 2.2e+02;
Matches 10; Conservative 6; Mismatches 14; Indels 0; Gaps 0;

QY      1 PIGNLNNVNGNHLIPAPPLPSQTDGAAR 30
Db      11 PVGRTGSDLEAVHQVQAPGLQFLAAGAG 40

RESULT 23
US-09-270-767-61121
; Sequence 61121, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
```

```
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 61121
; LENGTH: 55
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-61121

Query Match          30.9%; Score 51; DB 2; Length 55;
Best Local Similarity 53.3%; Pred. No. 18;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY      9 VNGNHLIPAPPLPS 23
Db      33 IGGGSVIPPFPVPS 47

RESULT 24
US-10-104-047-2072
; Sequence 2072, Application US/10104047
; Patent No. 6943241
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. 6943241el full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 2072
; LENGTH: 509
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-104-047-2072

Query Match          30.9%; Score 51; DB 2; Length 509;
Best Local Similarity 45.5%; Pred. No. 2.1e+02;
Matches 15; Conservative 4; Mismatches 10; Indels 4; Gaps 2;

QY      1 PIGNLNNVNGN-HLIPAPPL--PSQTDGAA 29
Db      406 PIGNLKEKELPDLQWCPAPLILILVQTERAA 438

RESULT 25
US-09-198-452A-63
; Sequence 63, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Griffiths, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragment,
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prev
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 63
; LENGTH: 644
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-198-452A-63

Query Match          30.9%; Score 51; DB 2; Length 644;
Best Local Similarity 50.0%; Pred. No. 2.8e+02;
Matches 9; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY      12 NHLIPAPPLPSQTDGAA 29
Db      11 PIGNLKEKELPDLQWCPAPLILILVQTERAA 438
```

Db 6 HHSYPPPPPPDOGVGAS 23

RESULT 26

US-09-080-897-4

; Sequence 4, Application US/09080897
; Patent No. 5985574

; GENERAL INFORMATION:

; APPLICANT: King, Mary-Claire

; APPLICANT: Lynch, Eric D.

; APPLICANT: Lee, Ming

; APPLICANT: Morrow, Jan E.

; APPLICANT: Welch, Piri L.

; APPLICANT: Leon, Pedro E.

; TITLE OF INVENTION: Modulators of Actin

; NUMBER OF SEQUENCES: 14

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP

; STREET: 75 DENISE DRIVE

; CITY: HILLSBOROUGH

; STATE: CALIFORNIA

; COUNTRY: USA

; ZIP: 94010

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/080,897

; FILING DATE:

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

; NAME: OSMAN, RICHARD A

; REGISTRATION NUMBER: 36,627

; REFERENCE/DOCKET NUMBER: UM97-001

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (650) 343-4341

; TELEFAX: (650) 343-4342

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1255 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; US-09-080-897-4

; Query Match 30.9%; Score 51; DB 1; Length 1255;

; Best Local Similarity 44.4%; Pred. No. 5.8e+02;

; Matches 8; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

; QY 8 NVNGNHLIPPAIPPLPSQT 25

; Db 564 SVSSSAVPPAPPLPGDS 581

; RESULT 27

; US-08-899-595-1

; Sequence 1, Application US/08899595

; Patent No. 6111072

; GENERAL INFORMATION:

; APPLICANT: Natumiya, Shun

; APPLICANT: Takahashi, No. 6111072uaki

; TITLE OF INVENTION: RHO TARGET PROTEIN HUMAN MDIA AND GENE

; NUMBER OF SEQUENCES: 14

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Foley & Lardner

; STREET: 3000 K Street, N.W., Suite 500

; CITY: Washington

; STATE: D.C.

; COUNTRY: USA

; ;

; ;

; ;

; ZIP: 20007-5109

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/899,595

; FILING DATE: 24-JUL-1997

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: JP 8-242701

; FILING DATE: 26-AUG-1996

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: JP 9-90170

; FILING DATE: 25-MAR-1997

; ATTORNEY/AGENT INFORMATION:

; NAME: Stephen A. Bent

; REGISTRATION NUMBER: 29,768

; REFERENCE/DOCKET NUMBER: 049441/0112

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (202)672-5300

; TELEFAX: (202)672-5399

; TELEX: 904136

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1255 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-08-899-595-1

; Query Match 30.9%; Score 51; DB 2; Length 1255;

; Best Local Similarity 44.4%; Pred. No. 5.8e+02;

; Matches 8; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

; QY 8 NVNGNHLIPPAIPPLPSQT 25

; Db 564 SVSSSAVPPAPPLPGDS 581

; RESULT 28

; US-09-323-735-4

; Sequence 4, Application US/09323735

; Patent No. 6197932

; GENERAL INFORMATION:

; APPLICANT: King, Mary-Claire

; APPLICANT: Lynch, Eric D.

; APPLICANT: Lee, Ming

; APPLICANT: Morrow, Jan E.

; APPLICANT: Welch, Piri L.

; APPLICANT: Leon, Pedro E.

; TITLE OF INVENTION: Modulators of Actin

; NUMBER OF SEQUENCES: 14

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP

; STREET: 75 DENISE DRIVE

; CITY: HILLSBOROUGH

; STATE: CALIFORNIA

; COUNTRY: USA

; ZIP: 94010

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/323,735

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 09/080,897

; FILING DATE:

; ;

; ;

```
ATTORNEY/AGENT INFORMATION:
NAME: OSMAN, RICHARD A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: UM97-001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 343-4341
TELEFAX: (650) 343-4342
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1255 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-323-735-4

Query Match
Best Local Similarity 30.6%; Score 51; DB 2; Length 1255;
Matches 8; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Qy 8 NNNGNHLPAPLPSTQ 25
Db 564 SVSSNAVPPAPLPEDS 581

RESULT 29
US-09-270-767-60585
Sequence 60585, Application US/09270767
Patent No. 6703491
GENERAL INFORMATION:
APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270,767
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 60585
LENGTH: 78
TYPE: PRT
ORGANISM: Drosophila melanogaster
US-09-270-767-60585

Query Match
Best Local Similarity 30.6%; Score 50.5; DB 2; Length 78;
Matches 13; Conservative 3; Mismatches 9; Indels 1; Gaps 1;

Qy 4 NLGNVNGNHLPPAPLPSTQDGA 29
Db 20 NLSMHNGGEL-PAATPLPAVSDEAA 44

RESULT 30
US-09-692-945-2
Sequence 2, Application US/09692945
Patent No. 6797695
GENERAL INFORMATION:
APPLICANT: Chiron Corporation
APPLICANT: Kyoto University
APPLICANT: Itoh, No. 6797695uyuki
APPLICANT: Kavanaugh, Michael W.
TITLE OF INVENTION: HUMAN RGF-20 GENE AND GENE EXPRESSION
FILE REFERENCE: 60219-6/16770.001
CURRENT APPLICATION NUMBER: US/09/692,945
CURRENT FILING DATE: 2000-10-20
NUMBER OF SEQ ID NOS: 17
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 212
TYPE: PRT
ORGANISM: Rattus norvegicus
US-09-692-945-2
```

```
Query Match
Best Local Similarity 30.6%; Score 50.5; DB 2; Length 212;
Matches 12; Conservative 3; Mismatches 9; Indels 3; Gaps 1;

Qy 5 LGNNVNGNHLPPA---PPLPSQTGA 28
Db 17 LGQGVSHFLPPAGRPPLGERRGA 43

RESULT 31
US-09-895-674A-1
Sequence 1, Application US/09895674A
Patent No. 6753311
GENERAL INFORMATION:
APPLICANT: Fertala, Andrzej
APPLICANT: Ko, Frank
TITLE OF INVENTION: Collagen and Collagen-like Peptide Containing Polymeric
FILE REFERENCE: DRE-0032
CURRENT APPLICATION NUMBER: US/09/895,674A
CURRENT FILING DATE: 2001-06-28
PRIOR APPLICATION NUMBER: PCT/US01/
PRIOR FILING DATE: 2001-06-25
PRIOR APPLICATION NUMBER: 60/ 214,034
NUMBER OF SEQ ID NOS: 1
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 234
TYPE: PRT
ORGANISM: Homo sapiens
US-09-895-674A-1

Query Match
Best Local Similarity 30.6%; Score 50.5; DB 2; Length 234;
Matches 12; Conservative 2; Mismatches 9; Indels 3; Gaps 1;

Qy 2 IGNNVNGNHLPPAPLPSTQDG 27
Db 3 VGPPSGNNGP---PPGPGSGKDG 25

RESULT 32
US-09-270-767-45089
Sequence 45089, Application US/09270767
Patent No. 6703491
GENERAL INFORMATION:
APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270,767
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 45089
LENGTH: 259
TYPE: PRT
ORGANISM: Drosophila melanogaster
US-09-270-767-45089

Query Match
Best Local Similarity 30.6%; Score 50.5; DB 2; Length 259;
Matches 13; Conservative 3; Mismatches 9; Indels 1; Gaps 1;

Qy 4 NLGNVNGNHLPPAPLPSTQDGA 29
Db 201 NLSMHNGGEL-PAATPLPAVSDEAA 225

RESULT 33
US-09-623-497-1
Sequence 1, Application US/09623497
```

```
; Patent No. 6706490
; GENERAL INFORMATION:
; APPLICANT: COOK, ANDREW
; APPLICANT: ROWLEY, MERRILL
; APPLICANT: MACKAY, IAN
; TITLE OF INVENTION: METHOD FOR THE DIAGNOSIS OF RHEUMATOID ARTHRITIS
; FILE REFERENCE: 017227/0167
; CURRENT APPLICATION NUMBER: US/09/623,497
; CURRENT FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: PCT/AU98/00176
; PRIOR FILING DATE: 1998-03-18
; PRIOR APPLICATION NUMBER: AU P05712/97
; PRIOR FILING DATE: 1997-03-19
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 347
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-623-497-1

Query Match          30.6%; Score 50.5; DB 2; Length 347;
Best Local Similarity 46.2%; Pred. No. 1.6e+02;
Matches 12; Conservative 2; Mismatches 9; Indels 3; Gaps 1;

QY      2  IGNLNNVNGNHLIPAPPLPSQTDG 27
Db      155 VGPGSGNGNPG---PPGPPGSGKDG 177

RESULT 34
US-10-332-795-12
; Sequence 12, Application US/10332795
; Patent No. 6946253
; GENERAL INFORMATION:
; APPLICANT: The University Court of the University of Glasgow
; APPLICANT: Clements, John Barklie
; APPLICANT: Maclean, Alasdair Roderick
; TITLE OF INVENTION: HERPES ZINC FINGER MOTIFS
; FILE REFERENCE: 9013-51
; CURRENT APPLICATION NUMBER: US/10/332,795
; CURRENT FILING DATE: 2003-01-10
; PRIOR APPLICATION NUMBER: GB 0016890.6
; PRIOR FILING DATE: 2000-07-11
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 744
; TYPE: PRT
; ORGANISM: Human cytomegalovirus
US-10-332-795-12

Query Match          30.6%; Score 50.5; DB 2; Length 744;
Best Local Similarity 33.3%; Pred. No. 3.7e+02;
Matches 10; Conservative 3; Mismatches 6; Indels 11; Gaps 1;

QY      8  NVNGNHLI-----PPAPPLPSQTD 26
Db      692 SANGNHSTTATQOOQPPPPPPVPOQED 721

RESULT 35
US-08-468-996-10
; Sequence 10, Application US/08468996
; Patent No. 6645504
; GENERAL INFORMATION:
; APPLICANT: Weiner, Howard
; APPLICANT: Miller, Ariel
; APPLICANT: Zheng, Zhengji
; APPLICANT: Ahmed, Al-Sabbagh
; TITLE OF INVENTION: BYSTANDER SUPPRESSION OF TYPE 1 DIABETES BY ORAL ADMINISTRATION OF
; FILE REFERENCE: 1010/16959-USA
```

```
; CURRENT APPLICATION NUMBER: US/08/468,996
; CURRENT FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: US 07/843,752
; PRIOR FILING DATE: 1992-02-28
; PRIOR APPLICATION NUMBER: US 07/460,852
; PRIOR FILING DATE: 1990-02-21
; PRIOR APPLICATION NUMBER: US 07/596,936
; PRIOR FILING DATE: 1990-10-15
; PRIOR APPLICATION NUMBER: US 07/065,734
; PRIOR FILING DATE: 1987-06-24
; PRIOR APPLICATION NUMBER: US 07/454,486
; PRIOR FILING DATE: 1989-12-20
; PRIOR APPLICATION NUMBER: US 07/487,732
; PRIOR FILING DATE: 1990-03-02
; PRIOR APPLICATION NUMBER: US 07/551,632
; PRIOR FILING DATE: 1990-07-10
; PRIOR APPLICATION NUMBER: US 07/379,778
; PRIOR FILING DATE: 1989-07-14
; PRIOR APPLICATION NUMBER: US 07/607,826
; PRIOR FILING DATE: 1990-10-31
; PRIOR APPLICATION NUMBER: US 07/595,468
; PRIOR FILING DATE: 1990-10-10
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 1017
; TYPE: PRT
; ORGANISM: Homo sapiens
US-08-468-996-10
```

```
Query Match          30.6%; Score 50.5; DB 2; Length 1017;
Best Local Similarity 46.2%; Pred. No. 5.3e+02;
Matches 12; Conservative 2; Mismatches 9; Indels 3; Gaps 1;
```

```
QY      2  IGNLNNVNGNHLIPAPPLPSQTDG 27
Db      705 VGPGSGNGNPG---PPGPPGSGKDG 727
```

```
RESULT 36
US-08-931-820-3
; Sequence 3, Application US/08931820
; Patent No. 6010863
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Assay for collagen degradation
; NUMBER OF SEQUENCES: 4
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/931,820
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 96202596.1
; FILING DATE:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1060 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEetical: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; TISSUE TYPE: Collagen type II
US-08-931-820-3
```

```
Query Match          30.6%; Score 50.5; DB 2; Length 1060;
```

Best Local Similarity 46.2%; Pred. No. 5.5e+02;
Matches 12; Conservative 2; Mismatches 9; Indels 3; Gaps 1;

QY 2 IGNLGNVNGNHLIPAPPLPSQTDG 27
DB 724 VGPPGSGNPG---PPGPPGPGSGKD 746

RESULT 37
US-08-963-825-20

; Sequence 20, Application US/08963825
; Patent No. 6110689

; GENERAL INFORMATION:

; APPLICANT: Oviest, Per

; APPLICANT: Bonde, Martin

; TITLE OF INVENTION: A Method for Assaying Collagen Fragments

; TITLE OF INVENTION: In Body Fluids, A Test Kit and Means for Carrying Out the

; TITLE OF INVENTION: Method and Use of the Method to Diagnose the Presence of

; NUMBER OF SEQUENCES: 21

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Darby & Darby PC

; STREET: 805 Third Avenue

; CITY: New York

; STATE: New York

; COUNTRY: USA

; ZIP: 10022

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/963,825

; FILING DATE:

; CLASSIFICATION: 436

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/08/187,319

; FILING DATE: 21-JAN-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: Gogoris, Adda C

; REGISTRATION NUMBER: 29,714

; REFERENCE/DOCKET NUMBER: 4305/08701

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 212-527-7700

; TELEFAX: 212-753-6237

; TELEX: 236687

; INFORMATION FOR SEQ ID NO: 20:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1418 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; ORIGINAL SOURCE:

; ORGANISM: Homo sapiens

; IMMEDIATE SOURCE:

; CLONE: COLLAGEN -ALPHA 1 (II)

US-08-963-825-20

Query Match 30.6%; Score 50.5; DB 2; Length 1418;

Best Local Similarity 46.2%; Pred. No. 7.6e+02;

Matches 12; Conservative 2; Mismatches 9; Indels 3; Gaps 1;

QY 2 IGNLGNVNGNHLIPAPPLPSQTDG 27

DB 836 VGPPGSGNPG---PPGPPGPGSGKD 858

RESULT 38

US-09-010-999-1

; Sequence 1, Application US/09010999

; Patent No. 6132976

; GENERAL INFORMATION:

; APPLICANT: Poole, Anthony R.

; APPLICANT: Hollander, Anthony P.

; TITLE OF INVENTION: IMMUNOASSAYS FOR THE MEASUREMENT OF

; TITLE OF INVENTION: COLLAGEN DENATURATION AND CLEAVAGE IN CARTILAGE

; NUMBER OF SEQUENCES: 16

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Foley & Lardner

; STREET: 3000 K Street, N.W., Suite 500

; CITY: Washington

; STATE: D.C.

; COUNTRY: USA

; ZIP: 20007-5109

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/010,999

; FILING DATE: 22-JAN-1998

; CLASSIFICATION: 4335

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/448,501

; FILING DATE: 17-JUL-1995

; APPLICATION NUMBER: US 07/984,123

; FILING DATE: 04-DEC-1992

; ATTORNEY/AGENT INFORMATION:

; NAME: Bent, Stephen A.

; REGISTRATION NUMBER: 29,768

; REFERENCE/DOCKET NUMBER: 032931/0212

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (202) 672-5300

; TELEFAX: (202) 672-5399

; TELEX: 904136

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1418 amino acids

; TYPE: amino acid

; STRANDEDNESS:

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; ORIGINAL SOURCE:

; ORGANISM: Human Type II Collagen

US-09-010-999-1

Query Match 30.6%; Score 50.5; DB 2; Length 1418;

Best Local Similarity 46.2%; Pred. No. 7.6e+02;

Matches 12; Conservative 2; Mismatches 9; Indels 3; Gaps 1;

QY 2 IGNLGNVNGNHLIPAPPLPSQTDG 27

DB 836 VGPPGSGNPG---PPGPPGPGSGKD 858

RESULT 39

US-09-500-811-20

; Sequence 20, Application US/09500811

; Patent No. 6323314

; GENERAL INFORMATION:

; APPLICANT: Oviest, Per

; APPLICANT: Bonde, Martin

; TITLE OF INVENTION: A Method for Assaying Collagen Fragments

; TITLE OF INVENTION: In Body Fluids, A Test Kit and Means for Carrying Out the

; TITLE OF INVENTION: Method and Use of the Method to Diagnose the Presence of

; NUMBER OF SEQUENCES: 21

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Darby & Darby PC

; STREET: 805 Third Avenue

; CITY: New York

; STATE: New York

COUNTRY: USA
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/500,811
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/187,319
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Gogorie, Adda C
REGISTRATION NUMBER: 29,714
REFERENCE/DOCKET NUMBER: 4305/08701
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-527-7700
TELEFAX: 212-753-6237
TELEX: 236687
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 1418 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
IMMEDIATE SOURCE:
CLONE: COLLAGEN -ALPHA 1 (II)
US-09-500-811-20

Query Match 30.6%; Score 50.5; DB 2; Length 1418;
Best Local Similarity 46.2%; Pred. No. 7.6e+02;
Matches 12; Conservative 2; Mismatches 9; Indels 3; Gaps 1;

QY 2 IGNLNNVNGNHLIPAPPLPSQTDG 27
DB 836 VGPPSGNPG--PPGPPSGKDG 858

RESULT 40
US-09-570-573-20
Sequence 20, Application US/09570573
Patent No. 6342361
GENERAL INFORMATION:
APPLICANT: Qvist, Per
APPLICANT: Bonde, Martin
TITLE OF INVENTION: A Method for Assaying Collagen Fragments
TITLE OF INVENTION: in Body Fluids, A Test Kit and Means for Carrying Out the
METHOD OF INVENTION: Method and Use of the Method to Diagnose the Presence of
TITLE OF INVENTION: Disorders Associated with the Metabolism of
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Darby & Darby PC
STREET: 805 Third Avenue
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/570,573
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/187,319

FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Gogorie, Adda C
REGISTRATION NUMBER: 29,714
REFERENCE/DOCKET NUMBER: 4305/08701
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-527-7700
TELEFAX: 212-753-6237
TELEX: 236687
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 1418 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
IMMEDIATE SOURCE:
CLONE: COLLAGEN -ALPHA 1 (II)
US-09-570-573-20

Query Match 30.6%; Score 50.5; DB 2; Length 1418;
Best Local Similarity 46.2%; Pred. No. 7.6e+02;
Matches 12; Conservative 2; Mismatches 9; Indels 3; Gaps 1;

QY 2 IGNLNNVNGNHLIPAPPLPSQTDG 27
DB 836 VGPPSGNPG--PPGPPSGKDG 858

Search completed: August 1, 2006, 21:48:16
Job time : 52 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: August 1, 2006, 21:58:45 ; Search time 182 Seconds
(without alignments)
76.354 Million cell updates/sec

Title: US-09-189-415D-7
Perfect score: 165
Sequence: 1 PIGNLGNNVNGNHLIPADPLPSQTDGAR 30

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2097797 seqs, 463214858 residues

Total number of hits satisfying chosen parameters: 2097797

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : Published Applications AA Main:*

- 1: /EMC_Celerra_SIDS3/Ptodata/2/pubpaa/US07_PUBCOMB.pep:*
- 2: /EMC_Celerra_SIDS3/Ptodata/2/pubpaa/US08_PUBCOMB.pep:*
- 3: /EMC_Celerra_SIDS3/Ptodata/2/pubpaa/US09_PUBCOMB.pep:*
- 4: /EMC_Celerra_SIDS3/Ptodata/2/pubpaa/US10_PUBCOMB.pep:*
- 5: /EMC_Celerra_SIDS3/Ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
- 6: /EMC_Celerra_SIDS3/Ptodata/2/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	123	74.5	558	US-11-134-563-12	Sequence 12, Appl
2	123	74.5	558	US-11-052-554A-20	Sequence 20, Appl
3	62	37.6	986	US-10-437-963-184318	Sequence 184318,
4	61	37.0	173	US-10-437-963-177853	Sequence 177853,
5	60	36.4	613	US-11-097-143-9273	Sequence 9273, Ap
6	59	35.8	146	US-10-424-599-187133	Sequence 187133,
7	59	35.8	329	US-10-714-887-116	Sequence 116, Ap
8	58.5	35.5	604	US-10-264-049-2829	Sequence 2829, Ap
9	58.5	35.5	987	US-10-312-352-13	Sequence 13, Appl
10	58.5	35.5	2486	US-10-745-237-260	Sequence 260, Appl
11	58	35.2	809	US-10-032-585-7691	Sequence 7691, Ap
12	58	35.2	1026	US-10-415-656-2	Sequence 2, Appl1
13	58	35.2	1567	US-09-835-232-2	Sequence 2, Appl1
14	57.5	34.8	1567	US-10-308-485-2	Sequence 2, Appl1
15	57.5	34.8	69	US-10-437-963-11409	Sequence 112409,
16	57	34.5	241	US-11-097-143-1062	Sequence 1062, Ap
17	56.5	34.2	22	US-09-879-957-67	Sequence 67, Appl
18	56.5	34.2	22	US-10-807-856-67	Sequence 67, Appl
19	56	33.9	42	US-10-626-832-157	Sequence 157, Appl
20	56	33.9	129	US-10-424-599-24972	Sequence 24972,
21	56	33.9	145	US-10-856-499-794	Sequence 794, Appl
22	56	33.9	187	US-10-437-963-189120	Sequence 189120,
23	56	33.9	410	US-10-214-766-33	Sequence 33, Appl
24	56	33.9	410	US-10-369-493-21572	Sequence 21572, A
25	56	33.9	3402	US-10-626-832-27	Sequence 27, Appl
26	55.5	33.6	167	US-10-424-599-229211	Sequence 229211,
27	55.5	33.6	301	US-10-739-930-9050	Sequence 9050, Ap

28	55.5	33.6	627	US-11-150-845-47	Sequence 47, Appl
29	55.5	33.6	675	US-10-606-060A-12	Sequence 12, Appl
30	55	33.3	223	US-10-425-115-291796	Sequence 291796,
31	55	33.3	31	US-10-739-930-10384	Sequence 10284, A
32	55	33.3	598	US-10-282-122A-64954	Sequence 64954, A
33	55	33.3	671	US-10-437-963-140066	Sequence 140066,
34	54.5	33.0	198	US-10-425-115-242739	Sequence 242739,
35	54.5	33.0	283	US-11-097-143-31545	Sequence 31545, A
36	54	32.7	98	US-10-437-963-180357	Sequence 180357,
37	54	32.7	202	US-10-767-701-59871	Sequence 59871, A
38	54	32.7	279	US-10-108-260A-4737	Sequence 4737, Ap
39	54	32.7	319	US-10-425-114-72305	Sequence 72305, A
40	54	32.7	359	US-10-424-599-167611	Sequence 167611,
41	54	32.7	367	US-10-425-114-72350	Sequence 72250, A
42	54	32.7	453	US-10-424-599-148508	Sequence 149508,
43	54	32.7	454	US-10-425-114-40125	Sequence 40125, A
44	54	32.7	473	US-11-097-143-20694	Sequence 20694, A
45	54	32.7	495	US-10-425-115-364526	Sequence 364526,
46	54	32.7	1238	US-11-097-143-10569	Sequence 10569, A
47	54	32.7	1238	US-10-480-330-4	Sequence 4, Appl1
48	54	32.7	1965	US-10-480-330-2	Sequence 2, Appl1
49	54	32.7	1966	US-10-480-330-6	Sequence 6, Appl1
50	54	32.7	1966	US-10-480-330-8	Sequence 8, Appl1
51	54	32.7	1966	US-10-480-330-10	Sequence 10, Appl
52	54	32.7	1966	US-10-480-330-12	Sequence 12, Appl
53	54	32.7	1966	US-10-480-330-14	Sequence 14, Appl
54	54	32.7	1966	US-10-480-330-16	Sequence 16, Appl
55	54	32.7	1966	US-10-480-330-18	Sequence 18, Appl
56	54	32.7	1966	US-10-480-330-20	Sequence 20, Appl
57	54	32.7	1966	US-10-480-330-22	Sequence 22, Appl
58	54	32.7	1966	US-10-480-330-24	Sequence 24, Appl
59	54	32.7	1966	US-10-480-330-26	Sequence 26, Appl
60	54	32.7	1966	US-10-480-330-28	Sequence 28, Appl
61	54	32.7	2468	US-10-246-330-4	Sequence 4, Appl1
62	54	32.7	2468	US-10-282-122A-66335	Sequence 66335, A
63	54	32.4	602	US-10-156-761-8377	Sequence 8377, Ap
64	53.5	32.4	70	US-10-425-115-195981	Sequence 171955,
65	53	32.1	201	US-10-437-963-171395	Sequence 171395,
66	53	32.1	226	US-11-097-143-8079	Sequence 8079, Ap
67	53	32.1	261	US-10-425-115-368010	Sequence 368010,
68	53	32.1	517	US-10-437-963-146045	Sequence 146045,
69	53	32.1	523	US-10-389-566-1464	Sequence 1464, Ap
70	53	32.1	581	US-10-374-780A-2556	Sequence 2556, Ap
71	53	32.1	581	US-10-389-566-1003	Sequence 1003, Ap
72	53	32.1	581	US-10-225-066A-272	Sequence 272, Appl
73	53	32.1	1553	US-10-437-963-107684	Sequence 107684,
74	53	32.1	1832	US-11-087-099-7762	Sequence 7762, Ap
75	53	32.1	1832	US-11-188-298-18195	Sequence 18195, A
76	53	32.1	1832	US-10-425-115-338456	Sequence 338456,
77	53	32.1	1832	US-11-097-143-30303	Sequence 30303, A
78	52.5	31.8	130	US-10-424-599-263767	Sequence 263767,
79	52.5	31.8	378	US-10-424-599-185745	Sequence 185745,
80	52.5	31.8	449	US-11-087-099-3606	Sequence 3606, Ap
81	52.5	31.8	491	US-10-425-115-299734	Sequence 299734,
82	52.5	31.8	55	US-10-425-115-325496	Sequence 325496,
83	52.5	31.8	94	US-10-425-115-320706	Sequence 320706,
84	52.5	31.8	94	US-10-425-115-320706	Sequence 320706,
85	52.5	31.8	94	US-10-425-115-320706	Sequence 320706,
86	52.5	31.8	94	US-10-425-115-320706	Sequence 320706,
87	52.5	31.8	94	US-10-425-115-320706	Sequence 320706,
88	52.5	31.8	94	US-10-425-115-320706	Sequence 320706,
89	52.5	31.8	94	US-10-425-115-320706	Sequence 320706,
90	52.5	31.8	94	US-10-425-115-320706	Sequence 320706,
91	52.5	31.8	94	US-10-425-115-320706	Sequence 320706,
92	52.5	31.8	94	US-10-425-115-320706	Sequence 320706,
93	52.5	31.8	94	US-10-425-115-320706	Sequence 320706,
94	52.5	31.8	94	US-10-425-115-320706	Sequence 320706,
95	52.5	31.8	94	US-10-425-115-320706	Sequence 320706,
96	52.5	31.8	94	US-10-425-115-320706	Sequence 320706,
97	52.5	31.8	94	US-10-425-115-320706	Sequence 320706,
98	52.5	31.8	94	US-10-425-115-320706	Sequence 320706,
99	52.5	31.8	94	US-10-425-115-320706	Sequence 320706,
100	52.5	31.8	94	US-10-425-115-320706	Sequence 320706,

ALIGNMENTS

RESULT 1
US-11-134-563-12
; Sequence 12, Application US/11134563
; Publication No. US20050287569A1
; GENERAL INFORMATION:
; APPLICANT: Leong, John M.
; APPLICANT: Campellone, Kenneth G.
; TITLE OF INVENTION: ESSEFNU NUCLEIC ACIDS AND PROTEINS AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: 07917-280001
; CURRENT APPLICATION NUMBER: US/11/134,563
; CURRENT FILING DATE: 2005-05-20
; PRIOR APPLICATION NUMBER: US 60/573,600
; PRIOR FILING DATE: 2004-05-20
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 12
; LENGTH: 558
; TYPE: PRT
; ORGANISM: Escherichia coli
US-11-134-563-12

Query Match 74.5%; Score 123; DB 6; Length 558;
Best Local Similarity 82.1%; Pred. No. 2; 7e-06;
Matches 23; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 PIGNGNVNGNHLIPAPPLPSQTDGA 28
|||||:|:|:|||||
DB 2 PIGNGNPNVNNSIPAPPLPSQTDGA 29

RESULT 2
US-11-052-554A-20
; Sequence 20, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; TITLE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO: 20
; LENGTH: 558
; TYPE: PRT
; ORGANISM: Escherichia coli 0157:H7
US-11-052-554A-20

Query Match 74.5%; Score 123; DB 6; Length 558;
Best Local Similarity 82.1%; Pred. No. 2; 7e-06;
Matches 23; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 PIGNGNVNGNHLIPAPPLPSQTDGA 28
|||||:|:|:|||||
DB 2 PIGNGNPNVNNSIPAPPLPSQTDGA 29

RESULT 3
US-10-437-963-184318
; Sequence 184318, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO: 184318
; LENGTH: 986
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_81321C.1.pep
US-10-437-963-184318

Query Match 37.6%; Score 62; DB 4; Length 986;
Best Local Similarity 70.6%; Pred. No. 98;
Matches 12; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 14 LIPAPPLPSQTDGAR 30
|||||:|:|:|||||
DB 133 LIPSPPLPGPTDMFAR 149

RESULT 4
US-10-437-963-177853
; Sequence 177853, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO: 177853
; LENGTH: 173
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1) --(173)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_75467C.1.pep
US-10-437-963-177853

Query Match 37.0%; Score 61; DB 4; Length 173;
Best Local Similarity 68.8%; Pred. No. 21;
Matches 11; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 10 NGNHLIPAPPLPSQT 25
|||:|:|:|:|:|
DB 144 NGXHGPPAPPPPGRT 159

RESULT 5
US-11-097-143-9273


```
; Sequence 9273, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; FILE REFERENCE: DROSOPHILA GENES.
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9273
; LENGTH: 613
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-11-097-143-9273
```

```
Query Match 36.4%; Score 60; DB 6; Length 613;
Best Local Similarity 47.8%; Pred. No. 1e+02;
Matches 11; Conservative 3; Mismatches 9; Indels 0; Gaps 0;
```

QY 2 IGTLGNVNGNHLIPAPPLPSQ 24

DB 449 IGMTGNQLNSLSLPPPPVPPDQ 471

```
RESULT 6
US-10-424-599-187133
; Sequence 187133, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 187133
; LENGTH: 146
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_139994C.1.Pep
US-10-424-599-187133
```

```
Query Match 35.8%; Score 59; DB 4; Length 146;
Best Local Similarity 50.0%; Pred. No. 30;
Matches 12; Conservative 1; Mismatches 11; Indels 0; Gaps 0;
```

QY 1 PIGNLGNVNGNHLIPAPPLPSQ 24

I: || | | | | | | | |

DB 90 PLINLHQNPNSHLPPPPQPSQQQ 113

```
RESULT 7
US-10-714-887-116
; Sequence 116, Application US/10714887
; Publication No. US20060015972A1
; GENERAL INFORMATION:
; APPLICANT: Mendel Biotechnology, Inc.
; APPLICANT: HEARD, Jacqueline
; APPLICANT: RIECHMANN, Jose Luis
; APPLICANT: CREELMAN, Robert
; APPLICANT: RATCLIFFE, Oliver
; APPLICANT: CANALES, Roger
; APPLICANT: REPERTI, Peter
; APPLICANT: KUMIMOTO, Roderick W
; APPLICANT: GUTTERSON, Neal
; APPLICANT: REUBER, T. Lynne
; APPLICANT: PINEDA, Omaira
; APPLICANT: SHERMAN, Bradley K
; TITLE OF INVENTION: PLANT TRANSCRIPTIONAL REGULATORS OF DROUGHT STRESS
; FILE REFERENCE: MB10058-CIP
; CURRENT APPLICATION NUMBER: US/10/714,887
; CURRENT FILING DATE: 2003-11-13
; PRIOR APPLICATION NUMBER: 10/412,699
; PRIOR FILING DATE: 2003-04-10
; PRIOR APPLICATION NUMBER: 09/506,720
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/135,134
; PRIOR FILING DATE: 1999-05-20
; PRIOR APPLICATION NUMBER: 09/394,519
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: 09/533,392
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,029
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/532,591
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,030
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/125,814
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: 09/713,994
; PRIOR FILING DATE: 2000-11-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 430
; SOFTWARE: Patent version 3.2
; SEQ ID NO 116
; LENGTH: 329
; TYPE: PRT
; ORGANISM: Oryza sativa (japonica cultivar-group)
; FEATURE:
; OTHER INFORMATION: G3683 polypeptide Orthologous to G2999
US-10-714-887-116
```

```
Query Match 35.8%; Score 59; DB 5; Length 329;
Best Local Similarity 38.5%; Pred. No. 70;
Matches 10; Conservative 4; Mismatches 12; Indels 0; Gaps 0;
```

QY 3 GNIGNVNGNHLIPAPPLPSQTDGA 28

DB 35 GGGGGGNGNRHNSPFPFPAAESKX 60

```
RESULT 8
US-10-264-049-2829
; Sequence 2829, Application US/10264049
; Publication No. US20040005579A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA133PI
; CURRENT APPLICATION NUMBER: US/10/264,049
```

```

; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/18569
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 2829
; LENGTH: 604
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (7)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (9)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (18)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; NAME/KEY: MISC FEATURE
; LOCATION: (22)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
US-10-264-049-2829
```

```

Query Match          35.5%; Score 58.5; DB 4; Length 604;
Best Local Similarity 44.4%; Pred. No. 1.5e+02;
Matches 12; Conservative 5; Mismatches 7; Indels 3; Gaps 1;
```

```
Oy      2  IGNLGNVNGNHLPPAPLPSPQTDGA 28
Db      529  VGGMPFSYGNAMIPVAPIP---DGA 552
```

```

RESULT 9
US-10-312-352-13
; Sequence 13, Application US/10312352
; Publication No. US20040053824A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.; TANG, Y. Tom
; APPLICANT: YUEN, Henry; AZIMZAI, Yalda
; APPLICANT: HE, Ann; BATRA, Saijeev
; APPLICANT: LO, Terence P.; NGUYEN, Damiel B.
; APPLICANT: BURKILL, John D.; MARCUS, Gregory A.
; APPLICANT: ZINGLER, Kurt A.; GANDHI, Ameena R.
; APPLICANT: LAL, Preeti G.; KEARNEY, Liam
; APPLICANT: BURFORD, Neil; YAO, Monique G.
; APPLICANT: CHAMLA, Narinder K.; ELLIOT, Vicki S.
; APPLICANT: ARVIZU, Chandra S.; KHAN, Farrah A.
; APPLICANT: BAUGHN, Marlan R.; HAFALIA, April, J.A.
; APPLICANT: POLICKY, Jennifer L.; AU-YOUNG, Janice K.
; APPLICANT: LU, Yan; BOROMSKY, Mark L.
; APPLICANT: YANG, Junming; GURURAJAN, Rajagopal
; APPLICANT: WARREN, Bridget A.; GIETZEN, Kimberly J.
; APPLICANT: XU, Yuning; KALLICK, Deborah A.
; APPLICANT: LEE, Ernestine A.; THANGAVEILU, Kavitha
; APPLICANT: DELEGEANE, Angelo M.; LEE, Sally
; TITLE OF INVENTION: EXTRACELLULAR MATRIX AND CELL ADHESION MOLECULES
; FILE REFERENCE: PF-0794 USN
; CURRENT APPLICATION NUMBER: US/10/312,352
; CURRENT FILING DATE: 2002-12-18
; PRIOR APPLICATION NUMBER: PCT/US01/21067
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/215,454
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 60/219,462
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: US 60/240,111
```

```

; PRIOR FILING DATE: 2000-10-12
; PRIOR APPLICATION NUMBER: US 60/240,106
; PRIOR FILING DATE: 2000-10-12
; PRIOR APPLICATION NUMBER: US 60/244,021
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: US 60/248,887
; PRIOR FILING DATE: 2000-11-14
; PRIOR APPLICATION NUMBER: US 60/249,570
; PRIOR FILING DATE: 2000-11-16
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PERL Program
; SEQ ID NO 13
; LENGTH: 987
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc-feature
; OTHER INFORMATION: Incyte ID No. US20040053824A1 5773251CD1
US-10-312-352-13
```

```

Query Match          35.5%; Score 58.5; DB 4; Length 987;
Best Local Similarity 44.4%; Pred. No. 2.6e+02;
Matches 12; Conservative 5; Mismatches 7; Indels 3; Gaps 1;
```

```
Oy      2  IGNLGNVNGNHLPPAPLPSPQTDGA 28
Db      912  VGGMPFSYGNAMIPVAPIP---DGA 935
```

```

RESULT 10
US-10-745-237-260
; Sequence 260, Application US/10745237
; Publication No. US2005027301A1
; GENERAL INFORMATION:
; APPLICANT: Cyclacel Limited
; APPLICANT: Glover, David
; APPLICANT: Bell, Graham
; APPLICANT: Frenz, Lisa
; APPLICANT: Midgley, Carol
; TITLE OF INVENTION: Cell Cycle Progression Proteins
; FILE REFERENCE: P015819MO CYK
; CURRENT APPLICATION NUMBER: US/10/745,237
; CURRENT FILING DATE: 2003-12-23
; PRIOR APPLICATION NUMBER: US 60/439,123
; PRIOR FILING DATE: 2003-01-10
; PRIOR APPLICATION NUMBER: US 60/468,402
; PRIOR FILING DATE: 2003-05-06
; NUMBER OF SEQ ID NOS: 600
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 260
; LENGTH: 2486
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: 075179; BAA31672
US-10-745-237-260
```

```

Query Match          35.5%; Score 58.5; DB 5; Length 2486;
Best Local Similarity 44.4%; Pred. No. 6.8e+02;
Matches 12; Conservative 5; Mismatches 7; Indels 3; Gaps 1;
```

```
Oy      2  IGNLGNVNGNHLPPAPLPSPQTDGA 28
Db      2411  VGGMPFSYGNAMIPVAPIP---DGA 2434
```

```

RESULT 11
US-10-032-585-7691
; Sequence 7691, Application US/10032585
; Publication No. US20030180953A1
; GENERAL INFORMATION:
; APPLICANT: Terry, Roemer D.
; APPLICANT: Bo, Jjiang
```

APPLICANT: Charles, Boone
APPLICANT: Howard, Bussey
TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
FILE REFERENCE: 10102-005-999
CURRENT APPLICATION NUMBER: US/10/032,585
CURRENT FILING DATE: 2001-12-20
NUMBER OF SEQ ID NOS: 8000
SOFTWARE: PatentIn version 3.1
SEQ ID NO 7691
LENGTH: 809
TYPE: PRT
ORGANISM: Candida albicans
US-10-032-585-7691

Query Match 35.2%; Score 58; DB 4; Length 809;
Best Local Similarity 47.6%; Pred. No. 2.4e+02;
Matches 10; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Qy 4 NLGNVNGNHLPPAPLP SQ 24
Db 17 NRHNSIGNMHLPPPPPPRQ 37

RESULT 12
US-10-415-656-2
Sequence 2, Application US/10415656
Publication No. US20050101773A1
GENERAL INFORMATION:
APPLICANT: Genix Bioscience GmbH
TITLE OF INVENTION: Genes required for viability and/or reproduction in C. elegans and
FILE REFERENCE: CE61823US
CURRENT APPLICATION NUMBER: US/10/415,656
CURRENT FILING DATE: 2003-09-08
PRIOR APPLICATION NUMBER: US 60/246,721
PRIOR FILING DATE: 2000-11-09
NUMBER OF SEQ ID NOS: 50
SOFTWARE: PatentIn version 3.1
SEQ ID NO 2
LENGTH: 1026
TYPE: PRT
ORGANISM: C. elegans
US-10-415-656-2

Query Match 35.2%; Score 58; DB 5; Length 1026;
Best Local Similarity 54.5%; Pred. No. 3.1e+02;
Matches 12; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

Qy 3 GNLGNVNGNHLPPAPLP SQ 24
Db 991 GNGVNPSSGQSGPPPPPPSQ 1012

RESULT 13
US-09-835-232-2
Sequence 2, Application US/09835232
Patent No. US20020098489A1
GENERAL INFORMATION:
APPLICANT: Leder, Philip
APPLICANT: Leder, Benjamin
TITLE OF INVENTION: FORMIN-2 NUCLEIC ACIDS AND POLYPEPTIDES
FILE REFERENCE: 00383/052002
CURRENT APPLICATION NUMBER: US/09/835,232
CURRENT FILING DATE: 2001-04-12
PRIOR APPLICATION NUMBER: US 60/196,811
PRIOR FILING DATE: 2000-04-13
NUMBER OF SEQ ID NOS: 22
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 1567
TYPE: PRT
ORGANISM: Mus musculus

US-09-835-232-2

Query Match 35.2%; Score 58; DB 3; Length 1567;
Best Local Similarity 57.9%; Pred. No. 4.8e+02;
Matches 11; Conservative 2; Mismatches 2; Indels 4; Gaps 1;

Qy 4 NLGNVNGNHLPPAPLP 22
Db 831 SFGNNCN---VPPAPLP 845

RESULT 14
US-10-308-485-2
Sequence 2, Application US/10308485
Publication No. US20030170683A1
GENERAL INFORMATION:
APPLICANT: Leder, Philip
APPLICANT: Leder, Benjamin
TITLE OF INVENTION: FORMIN-2 NUCLEIC ACIDS AND POLYPEPTIDES
FILE REFERENCE: 00383/052002
CURRENT APPLICATION NUMBER: US/10/308,485
CURRENT FILING DATE: 2002-12-03
PRIOR APPLICATION NUMBER: US/09/835,232
PRIOR FILING DATE: 2001-04-12
PRIOR APPLICATION NUMBER: US 60/196,811
PRIOR FILING DATE: 2000-04-13
NUMBER OF SEQ ID NOS: 22
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 1567
TYPE: PRT
ORGANISM: Mus musculus
US-10-308-485-2

Query Match 35.2%; Score 58; DB 4; Length 1567;
Best Local Similarity 57.9%; Pred. No. 4.8e+02;
Matches 11; Conservative 2; Mismatches 2; Indels 4; Gaps 1;

Qy 4 NLGNVNGNHLPPAPLP 22
Db 831 SFGNNCN---VPPAPLP 845

RESULT 15
US-10-437-963-112409
Sequence 112409, Application US/10437963
Publication No. US20040123343A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Bardazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 112409
LENGTH: 69
TYPE: PRT
ORGANISM: Oryza sativa
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_16297C.1.pcp
US-10-437-963-112409

Query Match 34.8%; Score 57.5; DB 4; Length 69;
Best Local Similarity 50.0%; Pred. No. 21;


```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/630,915
; FILING DATE: 03-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Mirock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-174
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 67:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Other
; OTHER INFORMATION: Biotinylated N-terminal
; SEQUENCE DESCRIPTION: SEQ ID NO: 67:
US-10-807-856-67

Query Match      34.2%; Score 56.5; DB 4; Length 22;
Best Local Similarity 50.0%; Pred. No. 8.1;
Matches 11; Conservative 3; Mismatches 5; Indels 3; Gaps 1;

QY      6 GNNVNGNHLIPAPPLPSQTDG 27
DB      4 GNYVNA---LPPGPPPLPAKNG 22

RESULT 19
US-10-626-832-157
; Sequence 157, Application US/10626832
; Publication No. US20050003342A1
; GENERAL INFORMATION:
; APPLICANT: Davis Poynter, Nick
; APPLICANT: Nugent, Josephine
; APPLICANT: Birch-Machin, Ian
; APPLICANT: Allen, George P
; TITLE OF INVENTION: Viral Marker
; FILE REFERENCE: 620-262
; CURRENT APPLICATION NUMBER: US/10/626,832
; CURRENT FILING DATE: 2003-07-25
; PRIOR APPLICATION NUMBER: US 60/398,576
; PRIOR FILING DATE: 2002-07-26
; NUMBER OF SEQ ID NOS: 259
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 157
; LENGTH: 42
; TYPE: PRT
; ORGANISM: Equine herpesvirus 1
US-10-626-832-157

Query Match      33.9%; Score 56; DB 5; Length 42;
Best Local Similarity 66.7%; Pred. No. 18;
Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY      15 IPPAPPLPSQTDGA 29
DB      3 LPPAPPLPOSTSKAA 17

RESULT 20
US-10-424-599-249972
; Sequence 249972, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Zhou Yihua
```

```

; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 249972
; LENGTH: 129
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_67753C.1.pep
US-10-424-599-249972

Query Match      33.9%; Score 56; DB 4; Length 129;
Best Local Similarity 44.0%; Pred. No. 60;
Matches 11; Conservative 3; Mismatches 11; Indels 0; Gaps 0;

QY      6 GNNVNGNHLIPAPPLPSQTDGA 30
DB      25 GRPVNMPWFPPEPPPLQANRAAK 49

RESULT 21
US-10-856-499-794
; Sequence 794, Application US/10856499
; Publication No. US20040259145A1
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; TITLE OF INVENTION: Modification of Gene Transcription
; FILE REFERENCE: 11000.1021C2
; CURRENT APPLICATION NUMBER: US/10/856,499
; CURRENT FILING DATE: 2004-05-28
; NUMBER OF SEQ ID NOS: 2370
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 794
; LENGTH: 145
; TYPE: PRT
; ORGANISM: Eucalyptus grandis
US-10-856-499-794

Query Match      33.9%; Score 56; DB 5; Length 145;
Best Local Similarity 38.5%; Pred. No. 68;
Matches 10; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

QY      3 GNLGNNVNGNHLIPAPPLPSQTDGA 28
DB      61 GGVGDNNNGYLDHSPLSVWPLKSDGS 86

RESULT 22
US-10-437-963-189120
; Sequence 189120, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
```

NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 189120
LENGTH: 187
TYPE: PRT
ORGANISM: Oryza sativa
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_85659C.1.dep
US-10-437-963-189120

Query Match 33.9%; Score 56; DB 4; Length 187;
Best Local Similarity 58.8%; Pred. No. 89;
Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 13 HLIAPPPLPSQTDGAA 29
| : | | | | : | | | |
Db 24 HHLPPAPPSPAGAGAA 40

RESULT 23
US-10-214-766-33
Sequence 33, Application US/10214766
Publication No. US20030084473A1

GENERAL INFORMATION:
APPLICANT: Gocal, Greg
TITLE OF INVENTION: NON-TRANSGENIC HERBICIDE RESISTANT PLANTS
FILE REFERENCE: CA1138
CURRENT APPLICATION NUMBER: US/10/214,766
CURRENT FILING DATE: 2002-08-09
PRIOR APPLICATION NUMBER: US 60/311,734
PRIOR FILING DATE: 2001-08-09
NUMBER OF SEQ ID NOS: 53
SOFTWARE: PatentIn version 3.2
SEQ ID NO 33
LENGTH: 410
TYPE: PRT
ORGANISM: Pyrococcus abyssi
US-10-214-766-33

Query Match 33.9%; Score 56; DB 4; Length 410;
Best Local Similarity 44.0%; Pred. No. 2e+02;
Matches 11; Conservative 2; Mismatches 12; Indels 0; Gaps 0;

Qy 2 IGNLGNVNGNHLIAPPAPLPSTQD 26
| : | | | | : | | | |
Db 62 ISKFGAQVNGKTIIPQELTPGKID 86

RESULT 24
US-10-369-493-21572
Sequence 21572, Application US/10369493
Publication No. US20030233675A1

GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xianfeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369,493
CURRENT FILING DATE: 2003-02-28
PRIOR APPLICATION NUMBER: US 60/360,039
PRIOR FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 21572
LENGTH: 410
TYPE: PRT
ORGANISM: Pyrococcus abyssi
US-10-369-493-21572

Query Match 33.9%; Score 56; DB 4; Length 410;
Best Local Similarity 44.0%; Pred. No. 2e+02;

Matches 11; Conservative 2; Mismatches 12; Indels 0; Gaps 0;
Qy 2 IGNLGNVNGNHLIAPPAPLPSTQD 26
| : | | | | : | | | |
Db 62 ISKFGAQVNGKTIIPQELTPGKID 86

RESULT 25
US-10-626-832-27

Sequence 27, Application US/10626832
Publication No. US2005003342A1
GENERAL INFORMATION:
APPLICANT: Davis Poynter, Nick
APPLICANT: Nugent, Josephine
APPLICANT: Birch-Machin, Ian
APPLICANT: Allen, George P
TITLE OF INVENTION: Viral Marker
FILE REFERENCE: 620-262
CURRENT APPLICATION NUMBER: US/10/626,832
CURRENT FILING DATE: 2003-07-25
PRIOR APPLICATION NUMBER: US 60/398,576
PRIOR FILING DATE: 2002-07-26
NUMBER OF SEQ ID NOS: 259
SOFTWARE: PatentIn version 3.1
SEQ ID NO 27
LENGTH: 3402
TYPE: PRT
ORGANISM: Equine herpesvirus 1
US-10-626-832-27

Query Match 33.9%; Score 56; DB 5; Length 3402;
Best Local Similarity 66.7%; Pred. No. 1.9e+03;
Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 15 IPPAPLPSTQDGA 29
: | | | | | | | | |
Db 2549 LPAPLPSTQSKAA 2563

RESULT 26
US-10-424-599-229211

Sequence 229211, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa Thomas J
APPLICANT: Kovacic David K
APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 229211
LENGTH: 167
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_49002C.1.dep
US-10-424-599-229211

Query Match 33.6%; Score 55.5; DB 4; Length 167;
Best Local Similarity 44.8%; Pred. No. 90;
Matches 13; Conservative 2; Mismatches 9; Indels 5; Gaps 1;

Qy 1 PIGNLGNVNGNHLI-----PPAPLPSTQ 24
| : | | | | : | | | |
Db 61 PTPNTTNNNNNLIQTNTNPSPPPPPQ 89

RESULT 27
US-10-739-930-9050

```
; Sequence 9050, Application US/10739930
; Publication No. US20040216190A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; FILE REFERENCE: 38-21(53377)B
; CURRENT APPLICATION NUMBER: US/10/739,930
; CURRENT FILING DATE: 2003-12-18
; NUMBER OF SEQ ID NOS: 11088
; SEQ ID NO 9050
; LENGTH: 301
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: GLYMA-23APR03-C34671_1.p
US-10-739-930-9050

Query Match          33.6%; Score 55.5; DB 5; Length 301;
Best Local Similarity 44.8%; Pred. No. 1.7e+02;
Matches 13; Conservative 2; Mismatches 9; Indels 5; Gaps 1;

QY      1 PIGNLNNGVNGNHLI-----PPAPLPSPQ 24
Db      61 PPNNTTNNNNNNNLIQNTNTPSPSPPPQ 89

RESULT 28
US-11-150-845-47
; Sequence 47, Application US/11150845
; Publication No. US20060003399A1
; GENERAL INFORMATION:
; APPLICANT: Cytokelectra, Inc.
; APPLICANT: Tomasevic, Nenad
; APPLICANT: Jia, Zhiheng
; APPLICANT: Sakowicz, Roman
; APPLICANT: Pierce, Daniel
; APPLICANT: Finer, Jeffrey
; TITLE OF INVENTION: HIGH THROUGHPUT ACTIN POLYMERIZATION ASSAY
; FILE REFERENCE: 020552-007720US
; CURRENT APPLICATION NUMBER: US/11/150,845
; CURRENT FILING DATE: 2005-06-10
; PRIOR APPLICATION NUMBER: US 60/673,444
; PRIOR FILING DATE: 2005-04-20
; PRIOR APPLICATION NUMBER: 60/578,949
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 47
; LENGTH: 627
; TYPE: PRT
; ORGANISM: Candida albicans
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1) .. (627)
; OTHER INFORMATION: FOR1 FH1-FH2 domain
US-11-150-845-47

Query Match          33.6%; Score 55.5; DB 6; Length 627;
Best Local Similarity 55.0%; Pred. No. 3.6e+02;
Matches 11; Conservative 4; Mismatches 4; Indels 1; Gaps 1;

QY      9 VNGN-HLPPAPPLPSQDGG 27
Db      120 LNSGSGVIPPAPPLPPSSG 139

RESULT 29
US-10-606-060A-12
; Sequence 12, Application US/10606060A
; Publication No. US20040058369A1
; GENERAL INFORMATION:
; APPLICANT: Syngenta
```

```
; APPLICANT: Jepson, Ian
; APPLICANT: Martinez, Alberto
; APPLICANT: Greenland, Andrew James
; TITLE OF INVENTION: A GENE SWITCH
; FILE REFERENCE: 1392/4/3/2
; CURRENT APPLICATION NUMBER: US/10/606,060A
; CURRENT FILING DATE: 2003-06-25
; PRIOR APPLICATION NUMBER: US 08/653,648
; PRIOR FILING DATE: 1996-05-24
; PRIOR APPLICATION NUMBER: US 09/564,418
; PRIOR FILING DATE: 2000-05-03
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 12
; LENGTH: 675
; TYPE: PRT
; ORGANISM: Aedes aegypti
US-10-606-060A-12

Query Match          33.6%; Score 55.5; DB 4; Length 675;
Best Local Similarity 52.0%; Pred. No. 3.9e+02;
Matches 13; Conservative 2; Mismatches 9; Indels 1; Gaps 1;

QY      2 IGNLNNGVNGNHLI-PPAPLPSPQ 25
Db      115 VGNLNGVNPNOTLIPLPSPQ 139

RESULT 30
US-10-425-115-291796
; Sequence 291796, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 291796
; LENGTH: 223
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_29205C.1.pep
US-10-425-115-291796

Query Match          33.3%; Score 55; DB 4; Length 223;
Best Local Similarity 56.2%; Pred. No. 1.4e+02;
Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY      15 IPPAPPLPSQDGAAR 30
Db      25 VPPPPVPGAPDAAR 40

RESULT 31
US-10-739-930-10284
; Sequence 10284, Application US/10739930
; Publication No. US20040216190A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; FILE REFERENCE: 38-21(53377)B
; CURRENT APPLICATION NUMBER: US/10/739,930
; CURRENT FILING DATE: 2003-12-18
; NUMBER OF SEQ ID NOS: 11088
; SEQ ID NO 10284
```

```

; LENGTH: 319
; TYPE: PRT
; ORGANISM: Tricicum aestivum
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(319)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: TRIAE-23APR03-C340_35.p
US-10-739-930-10284
```

```
Query Match          33.3%; Score 55; DB 5; Length 319;
Best Local Similarity 56.2%; Pred. No. 2.1e+02;
Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
```

```
Oy 15 IPPAPPLPSQTDGAAR 30
Db 25 VPPPPVPGADPAAR 40
```

RESULT 32

```
US-10-282-122A-64954
; Sequence 64954, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zykand, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: EITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See file Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 64954
; LENGTH: 598
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-282-122A-64954
```

```
Query Match          33.3%; Score 55; DB 4; Length 598;
Best Local Similarity 60.0%; Pred. No. 4e+02;
Matches 12; Conservative 1; Mismatches 7; Indels 0; Gaps 0;
```

```
Oy 7 NNVNGNHLIPAPPLPSQTD 26
Db 343 NNVPSPPIPPAPPPSGLD 362
```

RESULT 33

```
US-10-437-963-140066
; Sequence 140066, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 140066
; LENGTH: 671
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_4129C.1.pcp
US-10-437-963-140066
```

```
Query Match          33.3%; Score 55; DB 4; Length 671;
Best Local Similarity 36.8%; Pred. No. 4.5e+02;
Matches 14; Conservative 6; Mismatches 8; Indels 10; Gaps 2;
```

```
Oy 2 IGNT-----GNVNGNHLIPAPPLPSQTDGAAR 30
Db 484 LGNTVAPVPGGNDNNGN-APPDPYPMATNEAAK 520
```

RESULT 34

```
US-10-425-115-242739
; Sequence 242739, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 242739
; LENGTH: 198
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_152961C.1.pcp
US-10-425-115-242739
```

```
Query Match          33.0%; Score 54.5; DB 4; Length 198;
Best Local Similarity 75.0%; Pred. No. 1.4e+02;
Matches 12; Conservative 1; Mismatches 2; Indels 1; Gaps 1;
```

```
Oy 16 PPAPP-LPSQTDGAAR 30
Db 44 PPLPPHLPSQTDGASR 59
```



```
RESULT 35
US-11-097-143-31545
; Sequence 31545, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; PRIOR FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 31545
; LENGTH: 283
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-11-097-143-31545

Query Match      33.0%; Score 54.5; DB 6; Length 283;
Best Local Similarity 45.8%; Pred. No. 2.1e+02;
Matches 11; Conservative 3; Mismatches 3; Indels 7; Gaps 1;

QY      13 HLIPAPP-----LPSQTDGA 29
      ||:|||||
Db      166 HLPPAPPDYDATTTPAETGPA 189

RESULT 36
US-10-437-963-180357
; Sequence 180357, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Bardazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 180357
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_77732C.1.pep
US-10-437-963-180357
```

```
Query Match      32.7%; Score 54; DB 4; Length 98;
Best Local Similarity 40.6%; Pred. No. 78;
Matches 13; Conservative 5; Mismatches 10; Indels 4; Gaps 1;

QY      3 GNLGNVNGNHLIPAPPPLPS-----QTDGAAR 30
      ||:|||||
Db      56 GNGGRDEGNNLANPRTATIPSDDDDSGGGAR 87

RESULT 37
US-10-767-701-59871
; Sequence 59871, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 59871
; LENGTH: 202
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: 7536450.pep
US-10-767-701-59871

Query Match      32.7%; Score 54; DB 4; Length 202;
Best Local Similarity 69.2%; Pred. No. 1.7e+02;
Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      16 PPAPPPLPSQTDGA 28
      ||:|||||
Db      184 PPSPLPQYVDGA 196
```

```
RESULT 38
US-10-108-260A-4737
; Sequence 4737, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20040005560A1el full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4737
; LENGTH: 279
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-108-260A-4737

Query Match      32.7%; Score 54; DB 4; Length 279;
Best Local Similarity 47.6%; Pred. No. 2.3e+02;
Matches 10; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY      10 NGNHLIPAPPPLPSQTDGAAR 30
      ||:|||||
Db      95 SGKLSPVPVPRPRTQTASPAR 115

RESULT 39
US-10-425-114-72305
; Sequence 72305, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
```


GenCore Version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using bw model

Run on: August 1, 2006, 21:59:15 ; Search time 30 Seconds
(without alignments)
66.044 Million cell updates/sec

Title: US-09-189-415D-7
Perfect score: 165
Sequence: 1 PIGNLGNVNNGNHLLPPAPLPDSQDGAAR 30

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 232337 seqs, 66044171 residues

Total number of hits satisfying chosen parameters: 232337

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database :

Published Applications AA New:*

- 1: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 2: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 3: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 4: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 5: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 6: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 7: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
- 8: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US60_NEW_PUB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	62	37.6	980	6	US-10-449-902-38007
2	62	37.6	980	6	US-10-449-902-45343
3	62	37.6	980	6	US-10-449-902-52811
4	56	33.9	118	6	US-10-449-902-41370
5	56	33.9	159	7	US-11-174-307B-3920
6	56	33.9	3488	7	US-11-063-439-257
7	56	33.9	3491	7	US-11-063-439-228
8	56	33.9	3502	7	US-11-063-439-236
9	55	32.3	109	6	US-10-953-349-25635
10	55	32.3	319	6	US-10-953-349-29058
11	55	32.3	3504	7	US-11-063-439-119
12	55	32.3	3617	7	US-11-063-439-284
13	55	32.3	3700	7	US-11-063-439-290
14	54	32.7	277	6	US-10-953-349-24738
15	54	32.7	278	6	US-10-953-349-24737
16	54	32.7	279	6	US-11-293-697-4737
17	54	32.7	296	6	US-10-953-349-24736
18	53	32.1	75	6	US-10-953-349-26514
19	53	32.1	139	7	US-11-056-355B-61177
20	53	32.1	228	6	US-10-449-902-43015
21	53	32.1	427	7	US-11-056-355B-97135
22	53	32.1	452	7	US-11-056-355B-97134
23	53	32.1	517	6	US-10-449-902-48393
24	53	32.1	581	6	US-10-374-780A-2556

26	53	32.1	612	7	US-11-056-355B-97133	Sequence 97133, A
27	53	32.1	1481	7	US-11-056-355B-85230	Sequence 85230, A
28	53	32.1	1499	7	US-11-056-355B-78879	Sequence 78879, A
29	53	32.1	1510	7	US-11-056-355B-85229	Sequence 85229, A
30	53	32.1	1520	7	US-11-056-355B-85228	Sequence 85228, A
31	53	32.1	1528	7	US-11-056-355B-78878	Sequence 78878, A
32	53	32.1	1538	7	US-11-056-355B-78877	Sequence 78877, A
33	53	32.1	1568	7	US-11-056-355B-96434	Sequence 96434, A
34	53	32.1	1597	7	US-11-056-355B-96433	Sequence 96433, A
35	53	32.1	1607	7	US-11-056-355B-96432	Sequence 96432, A
36	53	32.1	1632	7	US-11-330-403-11014	Sequence 11014, A
37	52	31.5	89	6	US-10-953-349-31117	Sequence 31117, A
38	52	31.5	112	6	US-10-449-902-31139	Sequence 31139, A
39	52	31.5	788	6	US-10-449-902-54720	Sequence 54720, A
40	52	31.5	1012	6	US-10-519-342-3	Sequence 3, App1
41	51.5	31.2	271	7	US-11-056-355B-74097	Sequence 74097, A
42	51.5	31.2	358	7	US-11-056-355B-74096	Sequence 74096, A
43	51	30.9	386	6	US-10-374-780A-526	Sequence 526, App
44	51	30.9	451	6	US-10-449-902-42011	Sequence 42011, A
45	51	30.9	481	7	US-11-056-355B-45390	Sequence 45390, A
46	51	30.9	481	7	US-11-056-355B-49053	Sequence 49053, A
47	51	30.9	497	7	US-11-056-355B-45389	Sequence 45389, A
48	51	30.9	497	7	US-11-056-355B-49052	Sequence 49052, A
49	51	30.9	498	7	US-11-056-355B-45388	Sequence 45388, A
50	51	30.9	498	7	US-11-056-355B-49051	Sequence 49051, A
51	51	30.9	3502	7	US-11-063-439-160	Sequence 160, App
52	51	30.9	3575	7	US-11-063-439-279	Sequence 279, App
53	50.5	30.6	332	6	US-10-374-780A-2234	Sequence 2234, App
54	50.5	30.6	486	6	US-10-449-902-36900	Sequence 36900, A
55	50	30.3	128	6	US-11-293-697-2594	Sequence 2594, App
56	50	30.3	289	7	US-11-056-355B-20701	Sequence 20701, A
57	50	30.3	301	7	US-11-056-355B-65949	Sequence 65949, A
58	50	30.3	342	7	US-11-056-355B-20700	Sequence 20700, A
59	50	30.3	372	7	US-11-056-355B-80699	Sequence 80699, A
60	50	30.3	415	7	US-11-056-355B-87752	Sequence 87752, A
61	50	30.3	421	7	US-11-056-355B-87751	Sequence 87751, A
62	50	30.3	428	7	US-11-056-355B-70409	Sequence 70409, A
63	50	30.3	434	7	US-11-056-355B-70408	Sequence 70408, A
64	50	30.3	572	7	US-11-056-355B-87750	Sequence 87750, A
65	50	30.3	585	7	US-11-056-355B-70407	Sequence 70407, A
66	50	30.3	684	7	US-11-056-355B-77286	Sequence 77286, A
67	50	30.3	772	7	US-11-056-355B-77285	Sequence 77285, A
68	50	30.3	1386	6	US-10-449-902-41304	Sequence 41304, A
69	50	30.3	3518	7	US-11-063-439-287	Sequence 287, App
70	49.5	30.0	51	7	US-11-222-810-62	Sequence 62, App1
71	49.5	30.0	221	7	US-11-056-355B-2439	Sequence 2439, App
72	49.5	30.0	278	7	US-11-222-810-19	Sequence 19, App1
73	49.5	30.0	294	7	US-11-222-810-9	Sequence 9, App1
74	49.5	30.0	352	7	US-11-174-307B-1108	Sequence 1108, App
75	49.5	30.0	389	7	US-11-056-355B-15686	Sequence 15686, A
76	49.5	30.0	483	6	US-10-449-902-34684	Sequence 34684, A
77	49	29.7	322	6	US-10-449-902-35587	Sequence 35587, A
78	49	29.7	367	6	US-10-449-902-39806	Sequence 39806, A
79	49	29.7	369	7	US-11-056-355B-16332	Sequence 16332, A
80	49	29.7	372	7	US-11-056-355B-41356	Sequence 41356, A
81	49	29.7	411	6	US-10-449-902-54492	Sequence 54492, A
82	49	29.7	760	6	US-10-449-902-36863	Sequence 36863, A
83	49	29.7	1018	6	US-10-449-902-53312	Sequence 53312, A
84	49	29.7	3481	7	US-11-063-439-86	Sequence 86, App1
85	49	29.7	3623	7	US-11-063-439-282	Sequence 282, App
86	49	29.7	3974	7	US-11-063-439-276	Sequence 276, App
87	48.5	29.4	163	7	US-11-056-355B-57940	Sequence 57940, A
88	48.5	29.4	197	6	US-10-980-370-30	Sequence 30, App1
89	48.5	29.4	200	6	US-10-980-370-28	Sequence 28, App1
90	48.5	29.4	203	6	US-10-980-370-26	Sequence 26, App1
91	48.5	29.4	209	6	US-10-980-370-24	Sequence 24, App1
92	48.5	29.4	211	6	US-10-980-370-22	Sequence 22, App1
93	48.5	29.4	211	6	US-10-980-370-22	Sequence 22, App1
94	48.5	29.4	212	6	US-10-374-780A-7732	Sequence 7732, App
95	48.5	29.4	513	7	US-11-330-403-11981	Sequence 11981, A
96	48.5	29.4	515	7	US-11-330-403-11952	Sequence 11952, A
97	48.5	29.4	1220	6	US-10-449-902-41253	Sequence 41253, A
98	48	29.1	157	6	US-10-449-902-51888	Sequence 51888, A

99	48	29.1	250	7	US-11-056-355B-1663	Sequence 16663, A
100	48	29.1	252	6	US-10-449-902-34835	Sequence 34835, A

ALIGNMENTS

RESULT 1

```
US-10-449-902-38007
; Sequence 38007, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT CDNA AND USES THEREOF
; FILE REFERENCE: MOA-A02051-US
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 38007
; LENGTH: 980
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-449-902-38007
```

Query Match 37.6%; Score 62; DB 6; Length 980;
Best Local Similarity 70.6%; Pred. No. 10;
Matches 12; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 14 LIPAPPLPSQTDGAAR 30

DB 133 LIPSPPLPGPTDMFAR 149

RESULT 2

```
US-10-449-902-45343
; Sequence 45343, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT CDNA AND USES THEREOF
; FILE REFERENCE: MOA-A02051-US
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 45343
; LENGTH: 980
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-449-902-45343
```

Query Match 37.6%; Score 62; DB 6; Length 980;
Best Local Similarity 70.6%; Pred. No. 10;
Matches 12; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 14 LIPAPPLPSQTDGAAR 30

DB 133 LIPSPPLPGPTDMFAR 149

RESULT 3

```
US-10-449-902-52811
; Sequence 52811, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT CDNA AND USES THEREOF
; FILE REFERENCE: MOA-A02051-US
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 52811
; LENGTH: 980
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-449-902-52811
```

Query Match 37.6%; Score 62; DB 6; Length 980;
Best Local Similarity 70.6%; Pred. No. 10;
Matches 12; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 14 LIPAPPLPSQTDGAAR 30

DB 133 LIPSPPLPGPTDMFAR 149

RESULT 4

```
US-10-449-902-41370
; Sequence 41370, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT CDNA AND USES THEREOF
; FILE REFERENCE: MOA-A02051-US
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 41370
; LENGTH: 118
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-449-902-41370
```

Query Match 33.9%; Score 56; DB 6; Length 118;
Best Local Similarity 45.5%; Pred. No. 6.1;
Matches 10; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

OY 4 NLGNVNGNHLIPAPPLPSQT 25

DB 59 NISNNVNTNIRPPSKPTSTT 80

RESULT 5

```
US-11-174-307B-3920
; Sequence 3920, Application US/11174307B
```

```

? Publication No. US20060143729A1
? GENERAL INFORMATION:
? APPLICANT: ALEXANDROV, Nikolai
? APPLICANT: BROVER, Vyacheslav
? TITLE OF INVENTION: NUCLEOTIDE SEQUENCES AND POLYPEPTIDES ENCODED THEREBY
? FILE REFERENCE: 2750-1601PUS2
? CURRENT APPLICATION NUMBER: US/11/174,307B
? PRIOR APPLICATION NUMBER: 60/583,671
? PRIOR FILING DATE: 2004-06-30
? PRIOR APPLICATION NUMBER: 60/583,781
? PRIOR FILING DATE: 2004-06-30
? PRIOR APPLICATION NUMBER: 60/583,651
? PRIOR FILING DATE: 2004-06-30
? NUMBER OF SEQ ID NOS: 5544
? SEQ ID NO 3920
? LENGTH: 159
? TYPE: PRT
? ORGANISM: Zea mays
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION:
? OTHER INFORMATION: Pfam Name: Extensin_2; Pfam Description: Extensin-like region
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION:
? OTHER INFORMATION: Pfam Name: FH2; Pfam Description: Formin Homology 2 Domain
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION:
? OTHER INFORMATION: Pfam Name: Pyr_redox_2; Pfam Description: Pyridine
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION:
? OTHER INFORMATION: Pfam Name: LMBR1; Pfam Description: LMBR1-like membrane protein
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION:
? OTHER INFORMATION: Pfam Name: GATAase_2; Pfam Description: Glutamine
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION:
? OTHER INFORMATION: GI Number: 21992; NR Description: extensin [Volvox carterii]
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION:
? OTHER INFORMATION: GI Number: 50938941; NR Description: putative diaphanous
? OTHER INFORMATION: homologue [Oryza sativa (japonica cultivar-group)]
? OTHER INFORMATION: >gi|27817931|dbj|BAC55695.1| putative diaphanous homologue [Oryza
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION:
? OTHER INFORMATION: GI Number: 6523547; NR Description: hydroxyproline-rich
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION:
? OTHER INFORMATION: GI Number: 20420; NR Description: extensin [Prunus dulcis]
? OTHER INFORMATION: >gi|99861|pir||S20790 extensin - almond >gi|445616|trf||1909363A
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION:
? OTHER INFORMATION: GI Number: 50940371; NR Description: putative pherophorin-dz1
? OTHER INFORMATION: protein [Oryza sativa (japonica cultivar-group)]
? OTHER INFORMATION: >gi|42408241|dbj|BAD09398.1| putative pherophorin-dz1 protein
US-11-174-307B-3920
Query Match 33.9%; Score 56; DB 7; Length 159;
Best Local Similarity 57.1%; Pred. No. 8.3;
Matches 12; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
10 NGNHLIPAPPLPSOTDGAAR 30
|||||

```

```

Db          63  NGAHARQPPPPRRHQROGAR  83

RESULT 6
US-11-063-439-257
; Sequence 257, Application US/11063439
; Publication No. US20060147371A1
GENERAL INFORMATION:
APPLICANT: TUSZYNSKI, JACK A.
APPLICANT: GREENWALD, HOWARD J.
APPLICANT: CURRY, STEPHEN H.
APPLICANT: GOSS, KENDRICK
TITLE OF INVENTION: WATER-SOLUBLE COMPOUND
FILE REFERENCE: 1034312-000027
CURRENT APPLICATION NUMBER: US/11/063,439
CURRENT FILING DATE: 2005-02-23
PRIOR APPLICATION NUMBER: 10/878,905
PRIOR FILING DATE: 2004-06-28
PRIOR APPLICATION NUMBER: 10/923,615
PRIOR FILING DATE: 2004-08-20
PRIOR APPLICATION NUMBER: 60/516,134
PRIOR FILING DATE: 2003-10-31
PRIOR APPLICATION NUMBER: 10/808,618
PRIOR FILING DATE: 2004-03-24
PRIOR APPLICATION NUMBER: 10/867,517
PRIOR FILING DATE: 2004-06-14
NUMBER OF SEQ ID NOS: 418
SOFTWARE: PatentIn version 3.3
SEQ ID NO 228
LENGTH: 3491
TYPE: PRT
ORGANISM: Mycosphaerella pini
US-11-063-439-228

Query Match          33.9%; Score 56; DB 7; Length 3488;
Best Local Similarity 44.0%; Pred. No. 2e+02;
Matches 11; Conservative 3; Mismatches 11; Indels 0; Gaps 0;

Qy          4  NLGNVNGNHLIPAPPLPSQTQDA  28
| | | | | | | | | | | | | | | | | |
Db          1661  NNNNNNNNNNNPPPPPPSSSSSS  1685

RESULT 7
US-11-063-439-228
; Sequence 228, Application US/11063439
; Publication No. US20060147371A1
GENERAL INFORMATION:
APPLICANT: TUSZYNSKI, JACK A.
APPLICANT: GREENWALD, HOWARD J.
APPLICANT: CURRY, STEPHEN H.
APPLICANT: GOSS, KENDRICK
TITLE OF INVENTION: WATER-SOLUBLE COMPOUND
FILE REFERENCE: 1034312-000027
CURRENT APPLICATION NUMBER: US/11/063,439
CURRENT FILING DATE: 2005-02-23
PRIOR APPLICATION NUMBER: 10/878,905
PRIOR FILING DATE: 2004-06-28
PRIOR APPLICATION NUMBER: 10/923,615
PRIOR FILING DATE: 2004-08-20
PRIOR APPLICATION NUMBER: 60/516,134
PRIOR FILING DATE: 2003-10-31
PRIOR APPLICATION NUMBER: 10/808,618
PRIOR FILING DATE: 2004-03-24
PRIOR APPLICATION NUMBER: 10/867,517
PRIOR FILING DATE: 2004-06-14
NUMBER OF SEQ ID NOS: 418
SOFTWARE: PatentIn version 3.3
SEQ ID NO 228
LENGTH: 3491
TYPE: PRT
ORGANISM: Mycosphaerella pini
US-11-063-439-228

```

Query Match 33.9%; Score 56; DB 7; Length 3491;
Best Local Similarity 44.0%; Pred. No. 2e+02;
Matches 11; Conservative 3; Mismatches 11; Indels 0; Gaps 0;

QY 4 NLGNVNGNHLIPAPPLPSQTDGA 28
DB 1663 NNNNNNNNNNNPPPPPPSSSSS 1687

RESULT 8
US-11-063-439-236
; Sequence 236, Application US/11063439
; Publication No. US20060147371A1
; GENERAL INFORMATION:
; APPLICANT: TUSZYNSKI, JACK A.
; APPLICANT: GREENWALD, HOWARD J.
; APPLICANT: CURRY, STEPHEN H.
; APPLICANT: GOSK, KENDRICK
; TITLE OF INVENTION: WATER-SOLUBLE COMPOUND
; FILE REFERENCE: 1034312-000027
; CURRENT APPLICATION NUMBER: US/11/063,439
; CURRENT FILING DATE: 2005-02-23
; PRIOR APPLICATION NUMBER: 10/878,905
; PRIOR FILING DATE: 2004-06-28
; PRIOR APPLICATION NUMBER: 10/923,615
; PRIOR FILING DATE: 2004-08-20
; PRIOR APPLICATION NUMBER: 60/516,134
; PRIOR FILING DATE: 2003-10-31
; PRIOR APPLICATION NUMBER: 10/808,618
; PRIOR FILING DATE: 2004-03-24
; PRIOR APPLICATION NUMBER: 10/867,517
; PRIOR FILING DATE: 2004-06-14
; NUMBER OF SEQ ID NOS: 418
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 236
; LENGTH: 3502
; TYPE: PRT
; ORGANISM: Phaeosphaeria nodorum
US-11-063-439-236

Query Match 33.9%; Score 56; DB 7; Length 3502;
Best Local Similarity 44.0%; Pred. No. 2.1e+02;
Matches 11; Conservative 3; Mismatches 11; Indels 0; Gaps 0;

QY 4 NLGNVNGNHLIPAPPLPSQTDGA 28
DB 1670 NNNNNNNNNNNPPPPPPSSSSS 1694

RESULT 9
US-10-953-349-25635
; Sequence 25635, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 25635
; LENGTH: 109
; TYPE: PRT
; ORGANISM: Trifolium aestivum
US-10-953-349-25635

Query Match 33.3%; Score 55; DB 6; Length 109;
Best Local Similarity 56.2%; Pred. No. 7.4;
Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 15 IPPAPPLPSQTDGAAR 30
DB 46 VPPPPVPGAPDAAR 61

RESULT 10
US-10-953-349-29058
; Sequence 29058, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 29058
; LENGTH: 319
; TYPE: PRT
; ORGANISM: Trifolium aestivum
US-10-953-349-29058

Query Match 33.3%; Score 55; DB 6; Length 319;
Best Local Similarity 56.2%; Pred. No. 22;
Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 15 IPPAPPLPSQTDGAAR 30
DB 25 VPPPPVPGAPDAAR 40

RESULT 11
US-10-953-349-29057
; Sequence 29057, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 29057
; LENGTH: 346
; TYPE: PRT
; ORGANISM: Trifolium aestivum
US-10-953-349-29057

Query Match 33.3%; Score 55; DB 6; Length 346;
Best Local Similarity 56.2%; Pred. No. 24;
Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 15 IPPAPPLPSQTDGAAR 30
DB 52 VPPPPVPGAPDAAR 67

RESULT 12
US-11-063-439-119
; Sequence 119, Application US/11063439
; Publication No. US20060147371A1
; GENERAL INFORMATION:
; APPLICANT: TUSZYNSKI, JACK A.
; APPLICANT: GREENWALD, HOWARD J.
; APPLICANT: CURRY, STEPHEN H.
; APPLICANT: GOSK, KENDRICK
; TITLE OF INVENTION: WATER-SOLUBLE COMPOUND
; FILE REFERENCE: 1034312-000027
; CURRENT APPLICATION NUMBER: US/11/063,439

```
; CURRENT FILING DATE: 2005-02-23
; PRIOR APPLICATION NUMBER: 10/878,905
; PRIOR FILING DATE: 2004-06-28
; PRIOR APPLICATION NUMBER: 10/923,615
; PRIOR FILING DATE: 2004-08-20
; PRIOR APPLICATION NUMBER: 60/516,134
; PRIOR FILING DATE: 2003-10-31
; PRIOR APPLICATION NUMBER: 10/808,618
; PRIOR FILING DATE: 2004-03-24
; PRIOR APPLICATION NUMBER: 10/867,517
; PRIOR FILING DATE: 2004-06-14
; NUMBER OF SEQ ID NOS: 418
; SOFTWARE: Patentin version 3.3
; SEQ ID NO 119
; LENGTH: 3504
; TYPE: PRT
; ORGANISM: Galactomyces geotrichum
US-11-063-439-119
```

```
Query Match 33.3%; Score 55; DB 7; Length 3504;
Best Local Similarity 44.0%; Pred. No. 2.7e+02;
Matches 11; Conservative 3; Mismatches 11; Indels 0; Gaps 0;
```

```
Qy 5 LGNNVNGNHLIPAPPLPSQTDGAA 29
Db 224 INNNNNNNNPPPPPPSSSSSSAAA 248
```

```
RESULT 13
US-11-063-439-284
; Sequence 284, Application US/11063439
; Publication No. US20060147371A1
; GENERAL INFORMATION:
; APPLICANT: TUSZYNSKI, JACK A.
; APPLICANT: GREENWALD, HOWARD J.
; APPLICANT: CURRY, STEPHEN H.
; APPLICANT: GOSS, KENDRICK
; TITLE OF INVENTION: WATER-SOLUBLE COMPOUND
; FILE REFERENCE: 1034312-000027
; CURRENT APPLICATION NUMBER: US/11/063,439
; CURRENT FILING DATE: 2005-02-23
; PRIOR APPLICATION NUMBER: 10/878,905
; PRIOR FILING DATE: 2004-06-28
; PRIOR APPLICATION NUMBER: 10/923,615
; PRIOR FILING DATE: 2004-08-20
; PRIOR APPLICATION NUMBER: 60/516,134
; PRIOR FILING DATE: 2003-10-31
; PRIOR APPLICATION NUMBER: 10/808,618
; PRIOR FILING DATE: 2004-03-24
; PRIOR APPLICATION NUMBER: 10/867,517
; PRIOR FILING DATE: 2004-06-14
; NUMBER OF SEQ ID NOS: 418
; SOFTWARE: Patentin version 3.3
; SEQ ID NO 284
; LENGTH: 3617
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
US-11-063-439-284
```

```
Query Match 33.3%; Score 55; DB 7; Length 3617;
Best Local Similarity 50.0%; Pred. No. 2.8e+02;
Matches 11; Conservative 2; Mismatches 9; Indels 0; Gaps 0;
```

```
Qy 4 NIGNNVNGNHLIPAPPLPSQT 25
Db 1763 NNNNNNNNNNNPPPPPPPTT 1784
```

```
RESULT 14
US-11-063-439-290
; Sequence 290, Application US/11063439
; Publication No. US20060147371A1
; GENERAL INFORMATION:
```

```
; APPLICANT: TUSZYNSKI, JACK A.
; APPLICANT: GREENWALD, HOWARD J.
; APPLICANT: CURRY, STEPHEN H.
; APPLICANT: GOSS, KENDRICK
; TITLE OF INVENTION: WATER-SOLUBLE COMPOUND
; FILE REFERENCE: 1034312-000027
; CURRENT APPLICATION NUMBER: US/11/063,439
; CURRENT FILING DATE: 2005-02-23
; PRIOR APPLICATION NUMBER: 10/878,905
; PRIOR FILING DATE: 2004-06-28
; PRIOR APPLICATION NUMBER: 10/923,615
; PRIOR FILING DATE: 2004-08-20
; PRIOR APPLICATION NUMBER: 60/516,134
; PRIOR FILING DATE: 2003-10-31
; PRIOR APPLICATION NUMBER: 10/808,618
; PRIOR FILING DATE: 2004-03-24
; PRIOR APPLICATION NUMBER: 10/867,517
; PRIOR FILING DATE: 2004-06-14
; NUMBER OF SEQ ID NOS: 418
; SOFTWARE: Patentin version 3.3
; SEQ ID NO 290
; LENGTH: 3700
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-11-063-439-290
```

```
Query Match 33.3%; Score 55; DB 7; Length 3700;
Best Local Similarity 50.0%; Pred. No. 2.9e+02;
Matches 11; Conservative 2; Mismatches 9; Indels 0; Gaps 0;
```

```
Qy 4 NIGNNVNGNHLIPAPPLPSQT 25
Db 2408 NNNNNNNNNNNPPPPPPPTT 2429
```

```
RESULT 15
US-10-953-349-24738
; Sequence 24738, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: ENCODED THERBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: Patentin version 3.3
; SEQ ID NO 24738
; LENGTH: 277
; TYPE: PRT
; ORGANISM: Glycine max
US-10-953-349-24738
```

```
Query Match 32.7%; Score 54; DB 6; Length 277;
Best Local Similarity 43.5%; Pred. No. 26;
Matches 10; Conservative 4; Mismatches 9; Indels 0; Gaps 0;
```

```
Qy 7 NNNGNHLIPAPPLPSQTDGAA 29
Db 87 HRTNPPLAPLPPIPSPAVGS 109
```

```
RESULT 16
US-10-953-349-24737
; Sequence 24737, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: ENCODED THERBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
```

/ CURRENT FILING DATE: 2004-09-30
/ NUMBER OF SEQ ID NOS: 40252
/ SOFTWARE: PatentIn version 3.3
/ SEQ ID NO: 24737
/ LENGTH: 278
/ TYPE: PRT
/ ORGANISM: Glycine max
US-10-953-349-24737

Query Match 32.7%; Score 54; DB 6; Length 278;
Best Local Similarity 43.5%; Pred. No. 26;
Matches 10; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 7 NNVNGNHLPPAPLPSCQTDGA 29
DB 88 HRTNPNPLAPLPPIPSPAVGS 110

RESULT 17
US-11-293-697-4737
/ Sequence 4737, Application US/11293697
/ Publication No. US20060105376A1
/ GENERAL INFORMATION:
/ APPLICANT: HELIX RESEARCH INSTITUTE
/ TITLE OF INVENTION: Novel full length cDNA
/ FILE REFERENCE: H1-A0106
/ CURRENT APPLICATION NUMBER: US/11/293,697
/ PRIOR FILING DATE: 2005-12-05
/ PRIOR APPLICATION NUMBER: US/10/108,260
/ NUMBER OF SEQ ID NOS: 5458
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO: 4737
/ LENGTH: 279
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-11-293-697-4737

Query Match 32.7%; Score 54; DB 7; Length 279;
Best Local Similarity 47.6%; Pred. No. 26;
Matches 10; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 10 NGNHLPPAPLPSCQTDGA 30
DB 95 SGKLSPPVPPRPTGTASPAR 115

RESULT 18
US-10-953-349-24736
/ Sequence 24736, Application US/10953349
/ Publication No. US20060107345A1
/ GENERAL INFORMATION:
/ APPLICANT: ALEXANDROV, Nikolai et al.
/ TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
/ FILE REFERENCE: 2750-1579PUS2
/ CURRENT APPLICATION NUMBER: US/10/953,349
/ CURRENT FILING DATE: 2004-09-30
/ NUMBER OF SEQ ID NOS: 40252
/ SOFTWARE: PatentIn version 3.3
/ SEQ ID NO: 24736
/ LENGTH: 296
/ TYPE: PRT
/ ORGANISM: Glycine max
US-10-953-349-24736

Query Match 32.7%; Score 54; DB 6; Length 296;
Best Local Similarity 43.5%; Pred. No. 27;
Matches 10; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 7 NNVNGNHLPPAPLPSCQTDGA 29
DB 106 HRTNPNPLAPLPPIPSPAVGS 128

RESULT 19
US-10-953-349-26514
/ Sequence 26514, Application US/10953349
/ Publication No. US20060107345A1
/ GENERAL INFORMATION:
/ APPLICANT: ALEXANDROV, Nikolai et al.
/ TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
/ FILE REFERENCE: 2750-1579PUS2
/ CURRENT APPLICATION NUMBER: US/10/953,349
/ CURRENT FILING DATE: 2004-09-30
/ NUMBER OF SEQ ID NOS: 40252
/ SOFTWARE: PatentIn version 3.3
/ SEQ ID NO: 26514
/ LENGTH: 75
/ TYPE: PRT
/ ORGANISM: Triticum aestivum
US-10-953-349-26514

Query Match 32.1%; Score 53; DB 6; Length 75;
Best Local Similarity 81.8%; Pred. No. 8.7;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 16 PPAPLPSCQTD 26
DB 33 PPLPPLPQTD 43

RESULT 20
US-11-056-355B-61177
/ Sequence 61177, Application US/11056355B
/ Publication No. US20060150283A1
/ GENERAL INFORMATION:
/ APPLICANT: Brover, Vyacheslav
/ APPLICANT: Alexandrov, Nikolai
/ TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
/ FILE REFERENCE: 2750-1590PUS2
/ CURRENT APPLICATION NUMBER: US/11/056,355B
/ PRIOR FILING DATE: 2005-02-14
/ PRIOR APPLICATION NUMBER: 60/544,190
/ NUMBER OF SEQ ID NOS: 119966
/ SEQ ID NO: 61177
/ LENGTH: 139
/ TYPE: PRT
/ ORGANISM: Triticum aestivum
/ FEATURE:
/ NAME/KEY: peptide
/ LOCATION: (1)..(139)
/ OTHER INFORMATION: Ceres Seq. ID no. 15220132
US-11-056-355B-61177

Query Match 32.1%; Score 53; DB 7; Length 139;
Best Local Similarity 40.0%; Pred. No. 16;
Matches 10; Conservative 3; Mismatches 12; Indels 0; Gaps 0;

QY 6 GNNVNGNHLPPAPLPSCQTDGA 30
DB 8 GSRKKGNSLPPPPRRPAVDSAPR 32

RESULT 21
US-10-449-902-43015
/ Sequence 43015, Application US/10449902
/ Publication No. US20060123505A1
/ GENERAL INFORMATION:
/ APPLICANT: National Institute of Agrobiological Sciences.
/ APPLICANT: Bio-oriented Technology Research Advancement Institution.
/ APPLICANT: The Institute of Physical and Chemical Research.
/ APPLICANT: Foundation for Advancement of International Science.


```

; TITLE OF INVENTION: FULL-LENGTH PLANT CDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 43015
; LENGTH: 228
; TYPE: PRT
; ORGANISM: Oryza sativa
; US-10-449-902-43015

Query Match      32.1%; Score 53; DB 6; Length 228;
Best Local Similarity 52.6%; Pred. No. 28;
Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

Qy      11 GNHLPPAPLPSPQDGA 29
Db      199 GNKVPPPPPPAPSCNAEEA 217

RESULT 22
; US-11-056-355B-97135
; Sequence 97135, Application US/11056355B
; Publication No. US20060150283A1
; GENERAL INFORMATION:
; APPLICANT: Brover, Vyacheslav
; APPLICANT: Alexandrov, Nickolai
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
; FILE REFERENCE: 2750-1590PUS2
; CURRENT APPLICATION NUMBER: US/11/056,355B
; CURRENT FILING DATE: 2005-02-14
; PRIOR APPLICATION NUMBER: 60/544,190
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 119966
; SEQ ID NO 97135
; LENGTH: 427
; TYPE: prt
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
; LOCATION: (1)..(427)
; OTHER INFORMATION: Ceres Seq. ID no. 12736152
; US-11-056-355B-97135

Query Match      32.1%; Score 53; DB 7; Length 427;
Best Local Similarity 62.5%; Pred. No. 53;
Matches 10; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy      10 NGNHLPPAPLPSPQT 25
Db      54 NTNSLIPPPPPPPPOS 69

RESULT 23
; US-11-056-355B-97134
; Sequence 97134, Application US/11056355B
; Publication No. US20060150283A1
; GENERAL INFORMATION:
; APPLICANT: Brover, Vyacheslav
; APPLICANT: Alexandrov, Nickolai
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
; FILE REFERENCE: 2750-1590PUS2
; CURRENT APPLICATION NUMBER: US/11/056,355B
; CURRENT FILING DATE: 2005-02-14
; PRIOR APPLICATION NUMBER: 60/544,190
; PRIOR FILING DATE: 2004-02-13
```

```

; NUMBER OF SEQ ID NOS: 119966
; SEQ ID NO 97134
; LENGTH: 452
; TYPE: prt
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
; LOCATION: (1)..(452)
; OTHER INFORMATION: Ceres Seq. ID no. 12736151
; US-11-056-355B-97134

Query Match      32.1%; Score 53; DB 7; Length 452;
Best Local Similarity 62.5%; Pred. No. 56;
Matches 10; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy      10 NGNHLPPAPLPSPQT 25
Db      79 NTNSLIPPPPPPPPOS 94

RESULT 24
; US-10-449-902-48393
; Sequence 48393, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT CDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 48393
; LENGTH: 517
; TYPE: prt
; ORGANISM: Oryza sativa
; US-10-449-902-48393

Query Match      32.1%; Score 53; DB 6; Length 517;
Best Local Similarity 44.0%; Pred. No. 64;
Matches 11; Conservative 3; Mismatches 11; Indels 0; Gaps 0;

Qy      2 IGNLGNVNGNHLPPAPLPSPQTD 26
Db      65 VGNANSRVNGNSVQPTVELPSFKD 89

RESULT 25
; US-10-374-780A-2556
; Sequence 2556, Application US/10374780A
; Publication No. US20060162006A9
; GENERAL INFORMATION:
; APPLICANT: Sherman, Bradley K
; APPLICANT: Riechmann, Jose Luis
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Heard, Jacqueline E
; APPLICANT: Haake, Volker
; APPLICANT: Creelman, Robert A
; APPLICANT: Barcliffe, Oliver
; APPLICANT: Adam, Luc J
; APPLICANT: Reuber, T. Lynne
; APPLICANT: Keddie, James
; APPLICANT: Brown, Pierre E
; APPLICANT: Pilgrim, Mareha L
; APPLICANT: Dubell III, Arnold T
; APPLICANT: Pineda, Omaira
```

```

1  APPLICANT: Yu, Guo-liang
2  TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES IN PLANTS
3  FILE REFERENCE: MB1-0047 CIP
4  CURRENT FILING DATE: 2003-02-25
5  PRIOR APPLICATION NUMBER: 09/837,944
6  PRIOR FILING DATE: 2001-04-18
7  PRIOR APPLICATION NUMBER: 60/310,847
8  PRIOR FILING DATE: 2001-08-09
9  PRIOR APPLICATION NUMBER: 09/934,455
10 PRIOR FILING DATE: 2001-08-22
11 PRIOR APPLICATION NUMBER: 60/336,049
12 PRIOR FILING DATE: 2001-11-19
13 PRIOR APPLICATION NUMBER: 60/338,692
14 PRIOR FILING DATE: 2001-12-11
15 PRIOR APPLICATION NUMBER: 10/171,468
16 PRIOR FILING DATE: 2002-06-14
17 PRIOR APPLICATION NUMBER: 10/225,066
18 PRIOR FILING DATE: 2002-08-09
19 PRIOR APPLICATION NUMBER: 10/225,067
20 PRIOR FILING DATE: 2002-08-09
21 PRIOR APPLICATION NUMBER: 10/225,068
22 PRIOR FILING DATE: 2002-08-09
23 NUMBER OF SEQ ID NOS: 2906
24 SOFTWARE: PatentIn version 3.2
25 SEQ ID NO 2556
26 LENGTH: 581
27 TYPE: prt
28 ORGANISM: Arabidopsis thaliana
29 FEATURE:
30 OTHER INFORMATION: GI246
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
9
```

Query Match	32.1%	Score 53	DB 6	Length 581
Best local Similarity	47.6%	Pred. No. 73		
Match 10; Conservative	1	Mismatches	10	Indels 0
		Gaps	0	
Qy	1	P I G N I G N N V N G N H L I P P A P P L	21	
db	260	P R S Q I N N N N N G N F T R P P P L	280	

```

RESULT 26
US-11-056-355B-97133
: Sequence 97133, Application US/11056355B
: Publication No. US20060150283A1
: GENERAL INFORMATION:
: APPLICANT: Alexander, Vyacheslav
: APPLICANT: Brover, Vyacheslav
: TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
: FILE REFERENCE: 2750-1590PUS2
: CURRENT APPLICATION NUMBER: US/11/056,355B
: CURRENT FILING DATE: 2005-02-14
: PRIOR APPLICATION NUMBER: 60/544,190
: PRIOR FILING DATE: 2004-02-13
: NUMBER OF SEQ ID NOS: 119966
: SEQ ID NO 97133
: LENGTH: 612
: TYPE: prt
: ORGANISM: Arabidopsis thaliana
: FEATURE:
: NAME/KEY: peptide
: LOCATION: (1)..(612)
: OTHER INFORMATION: Ceres Seq. ID no. 12736150
US-11-056-355B-97133

```

Query Match	32.1%	Score 53	DB 7	Length 612
Best Local Similarity	62.5%	Pred. No. 77		
Matches	10	Conservative	5	Indels 0
				Gaps 0
0Y	10	NGNHLIPPAAPLSQT	25	

Db 239 NTNSLI PPPPPPPQS 254

```

RESULT 27
US-11-056-355B-85230
/ Sequence 85230, Application US/11056355B
/ Publication No. US2006015028A1
/ GENERAL INFORMATION:
/ APPLICANT: Brover, Vyacheslav
/ TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
/ TITLE OF INVENTION: Polypeptides Encoded Thereby
/ FILE REFERENCE: 2750-1590F052
/ CURRENT APPLICATION NUMBER: US/11/056,355B
/ CURRENT FILING DATE: 2005-02-14
/ PRIOR APPLICATION NUMBER: 60/544,190
/ PRIOR FILING DATE: 2004-02-13
/ NUMBER OF SEQ ID NOS: 119966
/ SEQ ID NO 85230
/ LENGTH: 1481
/ TYPE: prt
/ ORGANISM: Arabidopsis thaliana
/ FEATURE:
/ NAME/KEY: peptide
/ LOCATION: (1)..(1481)
/ OTHER INFORMATION: Ceres Seq. ID no. 12679560
US-11-056-355B-85230

```

Query Match	32.1%	Score 53	DB 7	Length 1481
Best Local Similarity	50.0%	Pred. No. 1.9e+02		
Matches 11	Conservative 1	Mismatches 10	Indels 0	Gaps 0
QY	3	GNLGNVNGNHLIPAPPLPSQ	24	
DB	454	GKSGRMKGYHLKVPPPLPQ	475	

```

RESULT 28
US-11-056-355B-78879
/ Sequence 78879, Application US/11056355B
/ Publication No. US20060150283A1
/ GENERAL INFORMATION:
/ APPLICANT: Alexandrov, Nicholai
/ APPLICANT: Brover, Vyacheslav
/ TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
/ TITLE OF INVENTION: Polypeptides Encoded Thereby
/ FILE REFERENCE: 2750-1590PUS2
/ CURRENT APPLICATION NUMBER: US/11/056,355B
/ CURRENT FILING DATE: 2005-02-14
/ PRIOR APPLICATION NUMBER: 60/544,190
/ PRIOR FILING DATE: 2004-02-13
/ NUMBER OF SEQ ID NOS: 119966
/ SEQ ID NO 78879
/ LENGTH: 1499
/ TYPE: prt
/ ORGANISM: Arabidopsis thaliana
/ FEATURE:
/ NAME/KEY: peptide
/ LOCATION: (1)..(1499)
/ OTHER INFORMATION: Ceres Seq. ID no. 12646479
US-11-056-355B-78879

```

Query Match	32.1%	Score 53	DB 7	Length 1499
Best Local Similarity	50.0%	Pred. No. 1.9e+02		
Matches 11	Conservative 1	Mismatches 10	Indels 0	Gaps 0
Qy	3	GNLGNVNGNHLIPAPPLPSQ	24	
	:			
Db	472	GKSGRMKGYHLKVPPLPQ	493	

RESULT 29
US-11-056-355B-85229

```
; Sequence 85229, Application US/11056355B
; Publication No. US20060150283A1
; GENERAL INFORMATION:
; APPLICANT: Brover, Vyacheslav
; APPLICANT: Alexandrov, Nickolai
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
; FILE REFERENCE: 2750-1590PUS2
; CURRENT APPLICATION NUMBER: US/11/056,355B
; PRIOR FILING DATE: 2005-02-14
; PRIOR APPLICATION NUMBER: 60/544,190
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 119966
; LENGTH: 1510
; TYPE: prt
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
; LOCATION: (1)..(1510)
; OTHER INFORMATION: Ceres Seq. ID no. 12679559
US-11-056-355B-85229

Query Match          32.1%; Score 53; DB 7; Length 1510;
Best Local Similarity 50.0%; Pred. No. 2e+02;
Matches 11; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

QY 3 GNLGNVNGNHLIPAPLPQ 24
DB 483 GKSGRWKGXHLKVPPPLPQ 504

RESULT 30
US-11-056-355B-85228
; Sequence 85228, Application US/11056355B
; Publication No. US20060150283A1
; GENERAL INFORMATION:
; APPLICANT: Brover, Vyacheslav
; APPLICANT: Alexandrov, Nickolai
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
; FILE REFERENCE: 2750-1590PUS2
; CURRENT APPLICATION NUMBER: US/11/056,355B
; PRIOR FILING DATE: 2005-02-14
; PRIOR APPLICATION NUMBER: 60/544,190
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 119966
; SEQ ID NO 85228
; LENGTH: 1520
; TYPE: prt
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
; LOCATION: (1)..(1520)
; OTHER INFORMATION: Ceres Seq. ID no. 12679558
US-11-056-355B-85228

Query Match          32.1%; Score 53; DB 7; Length 1520;
Best Local Similarity 50.0%; Pred. No. 2e+02;
Matches 11; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

QY 3 GNLGNVNGNHLIPAPLPQ 24
DB 493 GKSGRWKGXHLKVPPPLPQ 514

RESULT 31
US-11-056-355B-78878
; Sequence 78878, Application US/11056355B
; Publication No. US20060150283A1
; GENERAL INFORMATION:
; APPLICANT: Brover, Vyacheslav
; APPLICANT: Alexandrov, Nickolai
```

```
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
; FILE REFERENCE: 2750-1590PUS2
; CURRENT APPLICATION NUMBER: US/11/056,355B
; PRIOR FILING DATE: 2005-02-14
; PRIOR APPLICATION NUMBER: 60/544,190
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 119966
; SEQ ID NO 78878
; LENGTH: 1528
; TYPE: prt
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
; LOCATION: (1)..(1528)
; OTHER INFORMATION: Ceres Seq. ID no. 12646478
US-11-056-355B-78878

Query Match          32.1%; Score 53; DB 7; Length 1528;
Best Local Similarity 50.0%; Pred. No. 2e+02;
Matches 11; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

QY 3 GNLGNVNGNHLIPAPLPQ 24
DB 501 GKSGRWKGXHLKVPPPLPQ 522

RESULT 32
US-11-056-355B-78877
; Sequence 78877, Application US/11056355B
; Publication No. US20060150283A1
; GENERAL INFORMATION:
; APPLICANT: Brover, Vyacheslav
; APPLICANT: Alexandrov, Nickolai
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
; FILE REFERENCE: 2750-1590PUS2
; CURRENT APPLICATION NUMBER: US/11/056,355B
; PRIOR FILING DATE: 2005-02-14
; PRIOR APPLICATION NUMBER: 60/544,190
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 119966
; SEQ ID NO 78877
; LENGTH: 1538
; TYPE: prt
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
; LOCATION: (1)..(1538)
; OTHER INFORMATION: Ceres Seq. ID no. 12646477
US-11-056-355B-78877

Query Match          32.1%; Score 53; DB 7; Length 1538;
Best Local Similarity 50.0%; Pred. No. 2e+02;
Matches 11; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

QY 3 GNLGNVNGNHLIPAPLPQ 24
DB 511 GKSGRWKGXHLKVPPPLPQ 532

RESULT 33
US-11-056-355B-96434
; Sequence 96434, Application US/11056355B
; Publication No. US20060150283A1
; GENERAL INFORMATION:
; APPLICANT: Brover, Vyacheslav
; APPLICANT: Alexandrov, Nickolai
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
; FILE REFERENCE: 2750-1590PUS2
; CURRENT APPLICATION NUMBER: US/11/056,355B
; PRIOR FILING DATE: 2005-02-14
```

```
; PRIOR APPLICATION NUMBER: 60/544,190
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 119966
; SEQ ID NO 96434
; LENGTH: 1568
; TYPE: prt
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
; LOCATION: (1)..(1568)
; OTHER INFORMATION: Ceres Seq. ID no. 12733266
US-11-056-355B-96434

Query Match          32.1% Score 53; DB 7; Length 1568;
Best Local Similarity 50.0%; Pred. No. 2e+02;
Matches 11; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

Qy      3 GNLGNVNGNHLIPAPPLPSQ 24
Db      541 GKSGRMKGYHLKVPPPLPPQ 562

RESULT 34
US-11-056-355B-96433
; Sequence 96433, Application US/11056355B
; Publication No. US20060150283A1
; GENERAL INFORMATION:
; APPLICANT: Brover, Vyacheslav
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
; FILE REFERENCE: 2750-1590PUS2
; CURRENT APPLICATION NUMBER: US/11/056,355B
; PRIOR FILING DATE: 2005-02-14
; PRIOR APPLICATION NUMBER: 60/544,190
; NUMBER OF SEQ ID NOS: 119966
; SEQ ID NO 96433
; LENGTH: 1597
; TYPE: prt
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
; LOCATION: (1)..(1597)
; OTHER INFORMATION: Ceres Seq. ID no. 12733265
US-11-056-355B-96433

Query Match          32.1% Score 53; DB 7; Length 1597;
Best Local Similarity 50.0%; Pred. No. 2.1e+02;
Matches 11; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

Qy      3 GNLGNVNGNHLIPAPPLPSQ 24
Db      570 GKSGRMKGYHLKVPPPLPPQ 591

RESULT 35
US-11-056-355B-96432
; Sequence 96432, Application US/11056355B
; Publication No. US20060150283A1
; GENERAL INFORMATION:
; APPLICANT: Brover, Vyacheslav
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
; FILE REFERENCE: 2750-1590PUS2
; CURRENT APPLICATION NUMBER: US/11/056,355B
; PRIOR FILING DATE: 2005-02-14
; PRIOR APPLICATION NUMBER: 60/544,190
; NUMBER OF SEQ ID NOS: 119966
; SEQ ID NO 96432
; LENGTH: 1607
```

```
; TYPE: prt
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
; LOCATION: (1)..(1607)
; OTHER INFORMATION: Ceres Seq. ID no. 12733264
US-11-056-355B-96432

Query Match          32.1% Score 53; DB 7; Length 1607;
Best Local Similarity 50.0%; Pred. No. 2.1e+02;
Matches 11; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

Qy      3 GNLGNVNGNHLIPAPPLPSQ 24
Db      580 GKSGRMKGYHLKVPPPLPPQ 601

RESULT 36
US-11-330-403-11014
; Sequence 11014, Application US/11330403
; Publication No. US20060159563A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53629)B
; CURRENT APPLICATION NUMBER: US/11/330,403
; PRIOR FILING DATE: 2006-01-12
; NUMBER OF SEQ ID NOS: 19250
; SEQ ID NO 11014
; LENGTH: 1832
; TYPE: prt
; ORGANISM: Podospora anserina
US-11-330-403-11014

Query Match          32.1% Score 53; DB 7; Length 1832;
Best Local Similarity 45.5%; Pred. No. 2.4e+02;
Matches 10; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

Qy      1 PIGNLGNVNGNHLIPAPPLP 22
Db      1361 PIARGDENNGSASVPLPLP 1382

RESULT 37
US-10-953-349-31117
; Sequence 31117, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; PRIOR FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 31117
; LENGTH: 89
; TYPE: prt
; ORGANISM: Triticum aestivum
US-10-953-349-31117

Query Match          31.5% Score 52; DB 6; Length 89;
Best Local Similarity 57.9%; Pred. No. 14;
Matches 11; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy      6 GNVNNGNHLIPAPPLPSQ 24
Db      17 GNGVYANHEIIPPLPSLQ 35

RESULT 38
US-10-449-902-31399
```

```
; Sequence 3139, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agricultural Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 31399
; LENGTH: 112
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-449-902-31399

Query Match          31.5%; Score 52; DB 6; Length 112;
Best Local Similarity 56.2%; Pred. No. 17;
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 7 NNVNGNHLIPAPLP 22
   | : | | | | | : |
Db 62 NHPPGHGLIPMPPIP 77

RESULT 39
US-10-449-902-54720
; Sequence 54720, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agricultural Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 54720
; LENGTH: 788
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-449-902-54720

Query Match          31.5%; Score 52; DB 6; Length 788;
Best Local Similarity 57.1%; Pred. No. 13e+02;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 16 PPAPLPSCOTDGA 29
   | | | | | : | |
Db 272 PPPPPPPRTDNAS 285

RESULT 40
US-10-519-342-3
; Sequence 3, Application US/10519342
; Publication No. US20060160729A1
; GENERAL INFORMATION:
; APPLICANT: L4, Dean
; APPLICANT: Park, Kye Won
```

```
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR MANIPULATING THE GUIDED
; TITLE OF INVENTION: NAVIGATION OF ENDOTHELIAL TUBES DURING ANGIOGENESIS
; FILE REFERENCE: UTH-P01-010
; CURRENT APPLICATION NUMBER: US/10/519,342
; CURRENT FILING DATE: 2004-12-21
; PRIOR APPLICATION NUMBER: 60/392,142
; PRIOR FILING DATE: 2002-06-27
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1012
; TYPE: PRT
; ORGANISM: Mouse
US-10-519-342-3

Query Match          31.5%; Score 52; DB 6; Length 1012;
Best Local Similarity 44.4%; Pred. No. 1.7e+02;
Matches 12; Conservative 6; Mismatches 7; Indels 2; Gaps 2;

QY 3 GNLGNVNGNHLIPAP-PLPSQTDGA 28
   | : | | | | | | | | | : |
Db 856 GGVGSEV-GNLVYPPRCPPTPPSBS 881

Search completed: August 1, 2006, 22:02:45
Job time : 31 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using SW model

Run on: August 1, 2006, 21:41:55 ; Search time 39 Seconds
(without alignments)
74.013 Million cell updates/sec

Title: US-09-189-415D-7

Perfect score: 165

Sequence: 1 PIGNLGNVNGNHLIPAPPLPSQTDGAAR 30

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

1: PIR 80:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	123	74.5	558	2	A98199
2	123	74.5	558	2	E86045
3	58.5	35.5	968	2	T00353
4	58	35.2	1026	2	T20369
5	57	34.5	365	2	T24955
6	57	34.5	631	2	T32761
7	56.5	34.2	487	2	S65310
8	56	33.9	410	2	C75162
9	56	33.9	551	2	F83015
10	56	33.9	1420	2	T37781
11	56	33.9	3421	1	MZBBR6
12	55.5	33.6	624	2	S67301
13	55	33.3	442	2	T48170
14	55	33.3	687	2	T33077
15	55	33.3	1335	2	T18289
16	54.5	33.0	262	2	T48954
17	54	32.7	233	2	T28914
18	54	32.7	452	2	AES396
19	54	32.7	610	2	A83422
20	54	32.7	645	2	T16078
21	54	32.7	651	2	T15624
22	54	32.7	2468	2	A83412
23	53	32.1	193	2	T44106
24	53	32.1	569	2	B87180
25	53	32.1	612	2	T02414
26	53	32.1	770	2	T51024
27	52	31.5	276	2	AC0608
28	52	31.5	302	2	C86480
29	52	31.5	371	2	E88633

30	52	31.5	382	2	B88561	protein F58A4.7b [
31	52	31.5	406	2	T05865	hypothetical prote
32	52	31.5	554	2	T49833	related to Vea pro
33	52	31.5	725	2	T23722	hypothetical prote
34	52	31.5	725	2	T00492	hypothetical prote
35	51.5	31.2	164	2	T15525	hypothetical prote
36	51.5	31.2	186	2	T22685	hypothetical prote
37	51.5	31.2	274	2	PRMVM	HIV-1 retropepin
38	51.5	31.2	503	2	T48825	hypothetical prote
39	51.5	31.2	612	2	S65196	probable membrane
40	51	30.9	346	2	T46916	hypothetical prote
41	51	30.9	353	2	C70985	probable pxi1 pro
42	51	30.9	498	2	S20918	probable serine/th
43	51	30.9	543	2	S25128	61K protein - Auto
44	51	30.9	633	2	S62057	proline-rich prote
45	51	30.9	695	2	T24950	hypothetical prote
46	51	30.9	1255	2	T31065	diaphanous protein
47	51	30.9	1874	1	JU0533	genome polyprotein
48	50.5	30.6	212	2	JC7511	fibroblast growth
49	50.5	30.6	332	2	H84443	homodomain transc
50	50.5	30.6	744	1	QOBRN7	UL69 protein - hum
51	50.5	30.6	969	2	T15446	hypothetical prote
52	50.5	30.6	1418	2	T45467	collagen alpha 1(I
53	50.5	30.6	1487	1	CGH06C	collagen alpha 1(I
54	50	30.3	96	2	AF0134	hypothetical prote
55	50	30.3	100	2	T17126	hypothetical prote
56	50	30.3	168	2	S64830	hypothetical prote
57	50	30.3	204	2	S76259	hypothetical prote
58	50	30.3	220	2	C70525	hypothetical prote
59	50	30.3	342	2	A96511	unknown protein [I
60	50	30.3	382	2	E85082	hypothetical prote
61	50	30.3	382	2	T14186	hypothetical prote
62	50	30.3	440	2	AG1882	nitrate transport
63	50	30.3	516	2	H71332	hypothetical prote
64	50	30.3	517	2	E95959	probable membrane-
65	50	30.3	571	2	T43456	hypothetical prote
66	50	30.3	585	2	T00979	hypothetical prote
67	50	30.3	616	2	JC7905	fructan 6-fructosy
68	50	30.3	759	2	T00875	hypothetical prote
69	50	30.3	1171	2	T28701	probable polyketid
70	49.5	30.0	163	2	T51490	hydroxyproline-ric
71	49.5	30.0	202	2	JU0964	hypothetical prote
72	49.5	30.0	237	2	T35351	hypothetical prote
73	49.5	30.0	269	2	A83086	conserved hypotet
74	49.5	30.0	294	2	A55477	survival motor neu
75	49.5	30.0	448	2	T01570	hypothetical prote
76	49.5	30.0	1088	2	E86312	FL146.9 protein -
77	49.5	30.0	1325	2	T25753	hypothetical prote
78	49	29.7	105	2	T49751	hypothetical prote
79	49	29.7	114	2	T37080	hypothetical prote
80	49	29.7	289	2	T20177	hypothetical prote
81	49	29.7	302	2	T15936	hypothetical prote
82	49	29.7	316	2	T19291	hypothetical prote
83	49	29.7	316	2	T19288	hypothetical prote
84	49	29.7	409	2	S60975	hypothetical prote
85	49	29.7	442	1	A45390	gag polyprotein -
86	49	29.7	442	1	FOLJVS	gag polyprotein -
87	49	29.7	448	2	JU1161	Gag protein - Maed
88	49	29.7	464	2	JC7143	endoglucanase I -
89	49	29.7	577	2	T16333	hypothetical prote
90	49	29.7	691	2	T46476	hypothetical prote
91	49	29.7	695	2	S44049	nucleosapsid prote
92	49	29.7	831	2	T48442	hypothetical prote
93	49	29.7	845	2	T17291	hypothetical prote
94	49	29.7	1097	2	A56138	transcription fact
95	49	29.7	1217	2	T39427	probable myosin I
96	49	29.7	1387	2	JC5502	G-protein signalin
97	49	29.7	1456	2	T01397	LTR gag/pol polyp
98	48.5	29.4	142	2	S50662	hypothetical prote
99	48.5	29.4	211	2	JC7353	fibroblast growth
100	48.5	29.4	294	2	B70825	probable 3-hydroxy

ALIGNMENTS

RESULT 1

translocated intimin receptor Tir [imported] - Escherichia coli (strain O157:H7, substrain O157:H7, subsubstrain R1MD 0509952)
C/Species: Escherichia coli
C/Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C/Accession: A98199
R/Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.; gasawara, N.; Yasunaga, T.; Kunhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
A/Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gence
A/Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gence
A/Reference number: A99629; MUID:21156231; PMID:11285796
A/Accession: A98199
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-558 <HA>
A/Cross-references: UNIPROT:Q9R396; UNIPARC:UPI00000D00CA; GB:BA000007; PIDD:BA837984.1
A/Experimental source: strain O157:H7, subsubstrain R1MD 0509952
C/Genetics:
C/Gene: ECe4561

RESULT 2

```

Query Match 74.5%; Score 123; DB 2; Length 558;
Best Local Similarity 82.1%; Pred No. 3.5e-08;
Matches 23; Conservative 2; Mismatches 3; Indels 0; Gaps 0

```

RESULT 2

Es86045
 probable translocated intimin receptor protein tir [imported] - Escherichia coli (strain
 C:Species: Escherichia coli
 C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
 C:Accession: E86045
 R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
 Miller, L.; Grobeck, E.U.; Davis, N.W.; Jim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
 Nature 409, 529-533, 2001
 A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
 A:Reference number: AB5480; MUID:21074935; PMID:11206551
 A:Accession: E86045
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-558 <STO>
 A:Cross-references: UNIPROT:Q9R396; UNIPARC:UP100000D00CA; GB:AE005174; NID:g12518449; E
 A:Experimental source: strain O157:H7, substrain EDL933
 C:Genetics:
 A:Gene: tir

RESULT 3

Query	March	74.5%;	Score 123;	DB 2;	Length 556;
Best	Similarity	82.1%;	Pred. No. 3.5e-08;		
Local	23;	Conservative	2;	Mismatches 3;	Indels 0;
Matches				Gaps	0
Qy	1	PIGNIGNNVNGNHLLPPAPPLPSQTDGA	28		
db	2	PIGNIGNNPNVNSLPPAPPLPSQTDGA	29		

RESULT 3

hypothetical protein KIAA0697 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 09-Jul-2004
C:Accession: T00353
R: Ishikawa, K.; Nacase, T.; Suyama, M.; Miyajima, N.; Tanaka, A.; Kotani, H.; Nomura, N.
DNA Res. 5, 169-176, 1998
A:Title: Prediction of the coding sequences of unidentified human genes. X. The complete
A:Reference number: Z14142; MUID:9840380; PMID:9734811
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Accession: T00353

A;Molecule type: mRNA
A;Residues: 1-968 <ISH>
A;Cross-references: UNIPROT:O75179; UNIPARC:UPI000017C1C1; EMBL:AB014597; NID:g3327207;
A;Experimental source: brain, clone HK0486
C;Genetics:
A;Note: KIA0697

Query Match	35.5%	Score 58.5	DB 2,	Length 968;
Local Similarity	44.4%	Pred. No. 15,		
Matches	12;	Mismatches	7;	Gaps 1.
OY	2	IGNLGNNVNGNHLIPAPPLPSQTDGA	28	
	:	: - :		
Db	893	VGGMPFSYVGAMIPVPAPIP----	DGA 916	

RESULT 4

hypothetical protein D2045.1 - *Caenorhabditis elegans*
C|Species: *Caenorhabditis elegans*
C|Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C|Accession: J20369

A:Reference

A.Accession: J20369
A.Date: preliminary; translated from GB/EMBL/DBJ
A.Molecule type: DNA
A.Residues: 1-1026 <WTL>
A.Citdb-references: UNIPROT:Q18987, UNIPARC:UPI000004881F, EMBL:Z35639, PIDD:CAA84697.1;
A.Experimental source: clone D2045

A; Gene:

```

A:Map position:3
A.introns: 56/2; 88/3; 186/1; 273/2; 331/2; 410/1; 530/1; 604/1; 740/3; 845/3; 921/3; 92
Query Match          35.2%; Score 58; DB 2; Length 1026;
Best Local Similarity 54.5%; Pred. No. 18;
Matches 12; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY      3 GNLGNNVNGNHLIRPPAPPSQ 24
      |||||:|||||
Db      991 GNGGVNPSGGGSGPPPPPSQ 1012

```

RESULT 5

147352
hypothetical protein TIGI.8 - *Caenorhabditis elegans*
C|Species: *Caenorhabditis elegans*
C|Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C|Accession: T124955

A:Reference

A:Reference number: 213900
A:Accession: T24955
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-365 <WIL>
A:Cross-references: UNIPROT:Q9XUP5; UNIPARC:UPI0000076820; EMBL:Z81592; PIDD:CA804730.1

A;Gene:

A; Introns: 80/3; 161/2; 183/3; 213/2; 239/1

Query Match	34.5%	Score 57	DB 2	Length 365
Best Local Similarity	69.2%	Pred. No. 7.4		
Matches 9; Conservative			2; Indels	0; Gaps 0

Qy

Db 146 VPPAPMPITVDG 158

RESULT 6
T32761
hypothetical protein T12F5.5 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C:Accession: T32761
R:Clarke, K.; Wohlmann, P.; Rohlfing, T.; Bauer, C.
A:Description: The sequence of C. elegans cosmid T12F5.
A:Reference number: Z21221
A:Accession: T32761
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-631 <CAA>
A:Cross-references: UNIPARC:UPI0000002DE9; EMBL:AF039718; PIDN:AA896747.1; GSPDB:GN00019
C:Genetics:
A:Gene: CESP:T12F5.5
A:Map position: 1
A:introns: 54/2; 100/1; 133/3; 186/3; 556/2; 609/3

Query Match
Best Local Similarity 34.5%; Score 57; DB 2; Length 631;
Matches 10; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 6 GNNVNGNHLIPAPPLPSQ 24
DB 264 GNSLSTNPPMPVPPLPPQ 282

RESULT 7
S65310
probable membrane protein YPL277C - yeast (*Saccharomyces cerevisiae*)
N:Alternate names: hypothetical protein P0318
C:Species: *Saccharomyces cerevisiae*
C>Date: 10-Dec-1994 #sequence_revision 31-May-1996 #ext_change 09-Jul-2004
C:Accession: S65310; S65331
R:Duettehoeft, A.; Floeth, M.; Fritz, M.; Hilbert, H.; Moestl, D.
A:Submitted to the Protein Sequence Database, May 1996
A:Reference number: S64967
A:Accession: S65331
A:Molecule type: DNA
A:Residues: 1-487 <DUE>
A:Cross-references: UNIPROT:Q08989; UNIPARC:UPI000006C2C8; EMBL:Z73633; NID:G1370569; PI
A:Experimental source: strain S288C (AB972)
C:Genetics:
A:Cross-references: SGD:S0006198
A:Map position: 16L
C:Keywords: transmembrane protein
F:7-23/Domain: transmembrane #status predicted <TM1>
F:445-461/Domain: transmembrane #status predicted <TM2>

Query Match
Best Local Similarity 34.2%; Score 56.5; DB 2; Length 487;
Matches 12; Conservative 1; Mismatches 12; Indels 1; Gaps 1;

QY 3 GNLGNVNGNH-LIPAPPLPSQTDG 27
DB 146 GRFGNGTNGDHPKPPPPPPDEKG 171

RESULT 8
C75162
3-phosphoshikimate 1-carboxyvinyltransferase (arcs) PAB0306 - *Pyrococcus abyssi* (strain

C:Species: *Pyrococcus abyssi*
C>Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C:Accession: C75162
R:anonymous, Genoscope
A:Submitted to the EMBL Data Library, July 1999
A:Description: *Pyrococcus abyssi* genome sequence: insights into archaeal chromosome str
A:Reference number: A75001
A:Accession: C75162
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-410 <KAM>
A:Cross-references: UNIPROT:Q9V1H1; UNIPARC:UPI0000034469; GB:AJ248284; GB:AL096836; NIT
A:Experimental source: strain Orsay
C:Genetics:
A:Gene: arcs; PAB0306
C:Superfamily: 3-phosphoshikimate 1-carboxyvinyltransferase; 3-phosphoshikimate 1-carbo:
F:19-401/Domain: 3-phosphoshikimate 1-carboxyvinyltransferase homology <PSK>

Query Match
Best Local Similarity 33.9%; Score 56; DB 2; Length 410;
Matches 11; Conservative 2; Mismatches 12; Indels 0; Gaps 0;

QY 2 IGNLGNVNGNHLIPAPPLPSQTD 26
DB 62 ISKFGAQNNGNKRIPPOELTPGKID 86

RESULT 9
F83015
hypothetical protein PA5037 [imported] - *Pseudomonas aeruginosa* (strain PA01)
C:Species: *Pseudomonas aeruginosa*
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #ext_change 09-Jul-2004
C:Accession: F83015
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lam,
.; Lory, S.; Olson, M.V.
A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic path
A:Reference number: AB2950; M01D:20437337; PMID:10984043
A:Accession: F83015
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-551 <STO>
A:Cross-references: UNIPROT:Q9HND4; UNIPARC:UPI000000C5EDA; GB:AE004917; GB:AE004091; NIT
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA5037

Query Match
Best Local Similarity 33.9%; Score 56; DB 2; Length 551;
Matches 15; Conservative 3; Mismatches 11; Indels 10; Gaps 2;

QY 1 PIGNLGNVNGNHLIPAPPLPSQTDGAA 29
DB 370 PVTPLANNGVTPMHPVPAPEPTAPATPTPTGTTPARA 408

RESULT 10
T37781
probable cytoskeleton assembly control protein - fission yeast (*Schizosaccharomyces pombe*)
C:Species: *Schizosaccharomyces pombe*
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T37781
R:Oliver, K.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
A:Submitted to the EMBL Data Library, September 1995
A:Reference number: Z21746
A:Accession: T37781
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1420 <OLI>
A:Cross-references: UNIPROT:O13736; UNIPARC:UPI0000069639; EMBL:Z98529; PIDN:CAB11030.1;
A:Experimental source: strain 972h-, cosmid c16E8
C:Genetics:

A;introns: 286/1; 375/1; 406/1; 651/1

Query Match 33.3%; Score 55; DB 2; Length 687;

Best Local Similarity 60.0%; Pred. No. 28;

Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 12 NHIIPAPPPLPSQTD 26

Db 606 NRVIIPSPRIPRSTD 620

RESULT 15

T18289

racGAP protein - slime mold (Dictyostelium discoideum)

C;Species: Dictyostelium discoideum

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C;Accession: T18289

R;Ludbrook, S.B.; Eccleston, J.; Strom, M.

J. Biol. Chem. 272, 15682-15685, 1997

A;Title: Cloning of a rhogAP homolog from Dictyostelium discoideum.

A;Reference number: Z18858; MUID:97332648; PMID:9188459

A;Accession: T18289

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-1335 <LUD>

A;Cross-references: UNIPROT:O00886; UNIPARC:UPI0000077A50; EMBL:Y10159; NID:92190354; PI

C;Genetics:

A;Gene: racGAP

A;introns: 113/1; 205/3

Query Match 33.3%; Score 55; DB 2; Length 1335;

Best Local Similarity 50.0%; Pred. No. 60;

Matches 11; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

QY 1 PIGNGNVNGNHIIPAPPPLP 22

Db 589 PISNGSGSRNGSFLDSPPPPP 610

RESULT 16

T18954

hypothetical protein T15B3.140 - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 09-Jul-2004

C;Accession: T18954

R;Jordan, N.; Bengert, S.; Wiedemann, R.; Voss, H.; Unsel, M.; Mewes, H.W.; Rudd, S.;

submitted to the Protein Sequence Database, April 2000

A;Reference number: Z25009

A;Accession: T18954

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-262 <JOR>

A;Cross-references: UNIPROT:Q9LXW0; UNIPARC:UPI00000AAB06; EMBL:AL163975; GSPDB:GN00061;

C;Genetics:

A;Gene: ATSP-T15B3.140

A;Map position: 3

A;introns: 32/1; 158/3; 198/1; 226/3

Query Match 33.0%; Score 54.5; DB 2; Length 262;

Best Local Similarity 40.0%; Pred. No. 11;

Matches 12; Conservative 6; Mismatches 11; Indels 1; Gaps 1;

QY 2 IGNLG-NVNGNHIIPAPPPLPSQTDGAAR 30

Db 159 VGNVWMDSTNRVIAIPPPPLPSRRKGSXR 188

RESULT 17

T28914

hypothetical protein T26C11.1 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004

C;Accession: T28914

R;Martin, J.

submitted to the EMBL Data Library, December 1995

A;Description: The sequence of C. elegans cosmid T26C11.

A;Reference number: Z20542

A;Accession: T28914

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-233 <MAR>

A;Cross-references: UNIPROT:Q22813; UNIPARC:UPI000007B69B; EMBL:U41017; PIDN:AAC48210.1

A;Experimental source: strain Bristol N2; clone T26C11

C;Genetics:

A;Gene: CESP:T26C11.1

A;Map position: X

A;introns: 35/3; 53/2; 78/3; 126/3

Query Match 32.7%; Score 54; DB 2; Length 233;

Best Local Similarity 53.8%; Pred. No. 11;

Matches 14; Conservative 0; Mismatches 4; Indels 8; Gaps 2;

QY 5 LGNNVNGN-----LIIPAP---PLP 22

Db 145 LKNNTNKNHIDTASLIPIPIVQPLP 170

RESULT 18

AB3596

2-octaprenyl-6-methoxyphenol hydroxylase (EC 1.14.13.-) [imported] - Brucella melitensis;

C;Species: Brucella melitensis

C;Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004

C;Accession: AB3596

R;DeIacchio, V.G.; Kapatal, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Log, T.; Ivanova,

; Mazur, M.; Goldmann, E.; Selkov, E.; Elker, P.H.; Hagius, S.; O'Callaghan, D.; Letes

Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002

A;Title: The genome sequence of the facultative intracellular pathogen Brucella meliten

A;Reference number: AD3252; PMID:11756688

A;Accession: AB3596

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-452 <KUR>

A;Cross-references: UNIPROT:Q8YC38; UNIPROT:Q8FW87; UNIPARC:UPI00000585A5; GB:AE008918;

A;Experimental source: strain 16M

C;Genetics:

A;Gene: BMEI10694

A;Map position: II

C;Keywords: oxidoreductase

Query Match 32.7%; Score 54; DB 2; Length 452;

Best Local Similarity 52.6%; Pred. No. 23;

Matches 10; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 12 NHIIPAPPPLPSQTDGAAR 30

Db 433 SHLRLRLRLPGKADDAAR 451

RESULT 19

AB3422

hypothetical protein PA1797 [imported] - Pseudomonas aeruginosa (strain PA01)

C;Species: Pseudomonas aeruginosa

C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004

C;Accession: AB3422

R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br

adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kae, A.; Lardig, K.; Lam,

; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000

A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho

A;Reference number: AB3422

A;Accession: AB3422

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-610 <STO>

A;Cross-references: UNIPROT:Q912U5; UNIPARC:UPI000000C54A1; GB:AE004605; GB:AE004091; NID

A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA1797

Query Match 32.7%; Score 54; DB 2; Length 610;
Best Local Similarity 55.6%; Pred. No. 33;
Matches 10; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 10 NGNHLIPAPLPSPQTDG 27
DB 17 SGCHGLPPAPPSAAG 34

RESULT 20

116078
hypotheical protein F14D12.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 04-Mar-2000
C:Accession: T16078
R:Minx, P.

submitted to the EMBL Data Library, November 1995
A:Description: The sequence of C. elegans cosmid F14D12.
A:Reference number: Z18457

A:Accession: T16078
A:Status: Preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-645 <MIN>

A:Cross-references: UNIPARC:UPI00001793D0; EMBL:U41021; NID:g1086679; PID:g1086683; PIDN

C:Genetics:
A:Gene: CESP:F14D12.1
A:Introns: 5/3; 28/3; 113/3; 155/2; 226/1; 260/1; 298/2; 363/3; 462/3; 509/1; 567/2; 611
C:Superfamily: Caenorhabditis elegans hypotheical protein F14D12.1

Query Match 32.7%; Score 54; DB 2; Length 645;
Best Local Similarity 47.6%; Pred. No. 35;
Matches 10; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 4 NLGNVNGNHLIPAPLPSPQ 24
DB 235 SLPEYVNVSHVIRPPPEE 255

RESULT 21

115624
hypotheical protein C25H3.6 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T15624
R:Johnson, D.

submitted to the EMBL Data Library, June 1995
A:Description: The sequence of C. elegans cosmid C25H3.
A:Reference number: Z18379

A:Accession: T15624
A:Status: Preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-651 <JOH>

A:Cross-references: UNIPARC:UPI00001787C2; EMBL:U29535; NID:g868251; PID:g868256; PIDN:A

C:Genetics:
A:Gene: CESP:C25H3.6
A:Introns: 58/1; 213/3; 273/1; 446/2; 468/2; 484/2; 510/1; 552/1; 609/1

Query Match 32.7%; Score 54; DB 2; Length 651;
Best Local Similarity 44.0%; Pred. No. 35;
Matches 11; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

QY 2 IGNLGNVNGNHLIPAPLPSPQTD 26
DB 610 IENAGNEIVQVTLGPLSPITDQTE 634

RESULT 22

A83412

hypotheical protein PA1874 [imported] - Pseudomonas aeruginosa (strain PA01)

C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004

C:Accession: A83412
R:Stover, C.K.; Pham, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kass, A.; Lapid, K.; Lim,
L.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000

A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: A82950; MUID:20437337; PMID:10984043

A:Accession: A83412
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-2468 <STO>

A:Cross-references: UNIPROT:Q912M3; UNIPARC:UPI00000C54E3; GB:AE004613; GB:AE004091; NID
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA1874

Query Match 32.7%; Score 54; DB 2; Length 2468;
Best Local Similarity 44.0%; Pred. No. 1.7e+02;
Matches 11; Conservative 2; Mismatches 12; Indels 0; Gaps 0;

QY 1 PIGNLGNVNGNHLIPAPLPSPQ 25
DB 537 PIGQTTADANGMSFTSPSTPLDGT 561

RESULT 23

T44106
hypotheical protein [imported] - Staphylococcus aureus (fragment)

C:Species: Staphylococcus aureus
C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 11-May-2000
C:Accession: T44106
R:Itto, T.; Katayama, Y.; Hiramatsu, K.

Anticrib. Agents Chemother. 43, 1445-1458, 1999
A:Title: Cloning and nucleotide sequence determination of the entire mec DNA of pre-mech
A:Reference number: Z22733; MUID:99278010; PMID:10348769

A:Accession: T44106
A:Status: Preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-193 <ITO>

A:Cross-references: UNIPARC:UPI000017AC12; EMBL:D86934; PIDN:BA82208.1
A:Experimental source: strain N315

Query Match 32.1%; Score 53; DB 2; Length 193;
Best Local Similarity 50.0%; Pred. No. 12;
Matches 13; Conservative 2; Mismatches 9; Indels 2; Gaps 1;

QY 2 IGNLGNVNGNHLIPAPLPSPQ 25
DB 78 IGNLGNVNGNHLIPAPLPSPQ 103

RESULT 24

B87180
pyruvate (or indolepyruvate) decarboxylase [imported] - Mycobacterium leprae

C:Species: Mycobacterium leprae
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 05-Oct-2004
C:Accession: B87180
R:Cole, S.T.; Bigmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Ho

R.; Davies, R.M.; Devlin, K.; Duthey, S.; Feltwell, T.; Fraser, A.; Hamlin, N.; Holroyd,
eam, M.A.; Rutherford, K.M.
Nature 409, 1007-1011, 2001

A:Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; Sq
A:Title: Massive gene decay in the leprosy bacillus.
A:Reference number: A86909; MUID:21128732; PMID:11234002

A:Accession: B87180
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-569 <STO>

A:Cross-references: UNIPROT:Q9GBD6; UNIPARC:UPI00000C6E3D; GB:AL450380; NID:g13093715; P
C:Genetics:

A:Gene: pdc
C:Superfamily: pyruvate decarboxylase/indolepyruvate decarboxylase; thiamin pyrophosphat

Query Match 32.1%; Score 53; DB 2; Length 569;
Best Local Similarity 64.3%; Pred. No. 41;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 16 PPAPLPSPQTDGAA 29
DB 183 PPAPLPSPQTDGTS 196

RESULT 25
T02414
probable protein kinase [imported] - Arabidopsis thaliana
N:Alternate names: protein kinase homolog P27L4.5
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 05-Oct-2004
C:Accession: T02414; G84628
R:Rounsley, S.D.; Kaul, S.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, submitted to the EMBL Data Library, May 1998
A:Description: Arabidopsis thaliana chromosome II BAC F27L4 genomic sequence.
A:Reference number: Z14658
A:Accession: T02414
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-612 <R0U>
A:Cross-references: UNIPROT:O64825; UNIPARC:UPI0000048575; EMBL:AC004482; NID:g3152602; R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.L.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.; euser, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salberg, S.L.; Frazer, C.M.; Venter, J Nature 402, 761-768, 1999
A>Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: G84628
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-612 <STO>
A:Cross-references: UNIPARC:UPI0000048575; GB:AE002093; NID:g3152607; PIDN:AACT1086.1; C C:Genetics:
A:Gene: F27L4.5; Acc293770
A:Map position: 2

Query Match 32.1%; Score 53; DB 2; Length 612;
Best Local Similarity 62.5%; Pred. No. 44;
Matches 10; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 10 NGNHLIPAPLPSPQ 25
DB 239 NTNSLIPPPPPPS 254

RESULT 26
T51024
related to C2H2 zinc finger transcription factor D-5p1 [imported] - Neurospora crassa
N:Alternate names: protein B7F21.50
C:Species: Neurospora crassa
C>Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 09-Jul-2004
C:Accession: T51024
R:Schulte, U.; Aign, V.; Hohenseil, J.; Brandt, P.; Partmann, B.; Holland, R.; Nyakatura, submitted to the Protein Sequence Database, July 2000
A:Reference number: Z25286
A:Accession: T51024
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-770 <SCH>
A:Cross-references: UNIPROT:O9P319; UNIPARC:UPI000017B4B0; EMBL:AL389901; GSPDB:GN00116; C:Genetics:
A:Gene: NCSP:B7F21.50
A:Map position: 6
A:introns: 117/1

Query Match 32.1%; Score 53; DB 2; Length 770;
Best Local Similarity 40.0%; Pred. No. 58;
Matches 12; Conservative 3; Mismatches 5; Indels 10; Gaps 1;

Qy 6 GNNVNGNHLIP-----APPLPSPQ 25
DB 391 GNNNGSELASPDNPGYFGASPLPSSS 420

RESULT 27
AC0608
probable N-acetylmutamoyl-L-alanine amidase SRY0927 [imported] - Salmonella enterica sub
C:Species: Salmonella enterica subsp. enterica serovar Typh
A:Note: this species has also been called Salmonella typhi
C>Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C:Accession: AC0608
R:Fairhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K., A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A:Reference number: AB0502; MUID:21534947; PMID:11677608
A:Accession: AC0608
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-276 <PAR>
A:Cross-references: UNIPARC:UPI000005A0A0; GB:AL513382; PIDN:CAD05333.1; PID:g16502097; C:Genetics:
A:Gene: SRY0927

Query Match 31.5%; Score 52; DB 2; Length 276;
Best Local Similarity 56.2%; Pred. No. 24;
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 6 GNNVNGNHLIPAPLP 21
DB 65 GNNVSHYILPATPPL 80

RESULT 28
C86480
33.2K hypothetical protein - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: C86480
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K., anen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yi, G.; Frazer, C.M.; Venter, J.C.; Davis, R.W.
A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: C86480
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-302 <STO>
A:Cross-references: UNIPROT:O9C8B9; UNIPARC:UPI00000A8A47; GB:AE005172; NID:g11055830; F C:Genetics:
A:Map position: 1

Query Match 31.5%; Score 52; DB 2; Length 302;
Best Local Similarity 44.0%; Pred. No. 26;
Matches 11; Conservative 2; Mismatches 6; Indels 6; Gaps 1;

Qy 4 NLGNN-----VNGNHLIPAPLP 22
DB 63 NFVNNLISDDILNQTLLPPOPPP 87

```
RESULT 29
B88633
protein F56B3.1 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C:Accession: B88633
R:Anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog
A:Reference number: A75000; MUID:99069613; PMID:9851916
A:Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_elg
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A:Accession: B88633
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-371 <STO>
A:Cross-references: UNIPROT:O45114; UNIPARC:UPI00000772FC; GB:chr_IV; PIDN:AAC02612.1; F
C:Genetics:
A:Gene: P56B3.1
A:Map position: 4

Query Match
Best Local Similarity 31.5%; Score 52; DB 2; Length 371;
Matches 13; Conservative 1; Mismatches 14; Indels 0; Gaps 0;

OY 1 PIGNGNVNVGNHILPPAPPLPSQTDGA 28
Db 266 PPGNNGAPGAGGVGVPGPBPSPGSRGA 293

RESULT 30
B88561
protein F58A4.7b [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
C:Accession: B88561
R:Anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog
A:Reference number: A75000; MUID:99069613; PMID:9851916
A:Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_elg
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A:Accession: B88561
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-382 <STO>
A:Cross-references: UNIPARC:UPI000016415E; GB:chr_III; PIDN:CAA80170.1; PID:G3877829; GS
C:Genetics:
A:Gene: F58A4.7b
A:Map position: 3

Query Match
Best Local Similarity 31.5%; Score 52; DB 2; Length 382;
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

OY 8 NVNNGNHLPPAPPLPSQT 25
Db 361 HLEGSHFTPTSPPTSQRT 378

RESULT 31
T05865
hypothetical protein T29A15.90 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C:Accession: T05865
R:Beran, M.; Van Der Schueren, J.; Chung, Y.J.; Voet, M.; Robben, J.; Volckaert, G.; Ho
submitted to the Protein Sequence Database, March 1999
A:Reference number: Z15455
A:Accession: T05865
A:Molecule type: DNA
```

```
A:Residues: 1-406 <BEV>
A:Cross-references: UNIPROT:Q9T084; UNIPARC:UPI00000489FD; EMBL:AL035602; GSPDB:GN00062;
A:Experimental source: cultivar Columbia; BAC clone T29A15
C:Genetics:
A:Gene: ATSP:T29A15.90
A:Map position: 4
A:introns: 65/3; 153/2; 192/3; 262/2

Query Match
Best Local Similarity 31.5%; Score 52; DB 2; Length 406;
Matches 11; Conservative 9; Mismatches 8; Indels 8; Gaps 1;

OY 1 PIGNGNVNVGNHILPPAPPLPSQTDGA 28
Db 303 PFVSVTDGINSYGVGEMAIYIPSPCVPTDTGCA 338

RESULT 32
T49833
related to Vea protein [imported] - Neurospora crassa
N:Alternate names: protein B24H17.190
C:Species: Neurospora crassa
C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 09-Jul-2004
C:Accession: T49833
R:Schulte, U.; Aign, V.; Hohelsel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,
submitted to the Protein Sequence Database, May 2000
A:Reference number: Z25022
A:Accession: T49833
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-554 <SCH>
A:Cross-references: UNIPROT:Q9P526; UNIPARC:UPI00000695C4; EMBL:AL356815; GSPDB:GN00116;
A:Experimental source: BAC clone B24H17; strain OR74A
C:Genetics:
A:Gene: NCSP:B24H17.190
A:Map position: 6
A:introns: 56/1

Query Match
Best Local Similarity 31.5%; Score 52; DB 2; Length 554;
Matches 11; Conservative 5; Mismatches 13; Indels 2; Gaps 1;

OY 2 IGNLGS--NNVNGNHLPPAPPLPSQTDGAAR 30
Db 470 VSNNGGLTSAGGVNQLPPPPPPPVVAGSKR 500

RESULT 33
T23722
hypothetical protein M04G12.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T23722
R:Sim, M.
submitted to the EMBL Data Library, October 1996
A:Reference number: Z19789
A:Accession: T23722
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-561 <WIL>
A:Cross-references: UNIPROT:P92007; UNIPARC:UPI000016426E; EMBL:Z81103; PIDN:CAB03211.1;
A:Experimental source: clone M04G12
C:Genetics:
A:Gene: CESP:M04G12.4
A:Map position: 5
A:introns: 193/3; 458/1; 505/3; 553/2

Query Match
Best Local Similarity 31.5%; Score 52; DB 2; Length 561;
Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

OY 4 NIGNVNVGNHILPPAPPLP 22
::| |||: ||| |
```

Db 532 SMGYSNGNFPDP PPP 550

RESULT 34

T00492

hypothetical protein R29144_1 - human

C:Species: Homo sapiens (man)

C>Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 09-Jul-2004

C:Accession: T00492

R:landerin, J.E.; McCreedy, P.M.; Skowronski, E.; Adamson, A.W.; Burkhardt-Schultz, K.; G.

Christensen, M.; Georgescu, A.; Avila, J.; Liu, S.; Attix, C.; Andreise, T.; Trankler,

Arellano, A.; Montgomery, M.; Ow, D.; Nolan, M.; Trong, S.; Kobayashi, A.; Olsen, A.O.;

submitted to the EMBL Data Library, February 1998

A:Description: Sequence analysis of a 3.5 Mb contig in human 19p13.3 containing a serine

A:Reference number: Z14157

A:Accession: T00492

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-725 <LAMB>

A:Cross-references: UNIPROT:O43385; UNIPARC:UPI000006DF67; EMBL:AC004221; NID:g2911257;

C:Genetics:

A:Introns: 2/3; 76/2; 98/1; 144/1; 229/3; 276/3; 347/1; 405/2

C:Superfamily: human herpesvirus 4 BHLF1 protein

Query Match 31.2%; Score 51.5; DB 2; Length 725;
Best Local Similarity 60.0%; Pred. No. 73;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 16 PPAPLPSPQTDGAAR 30
Db 502 PPSPAPRPKDGSEAR 516

RESULT 35

T15525

hypothetical protein C16B8.3 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 05-Oct-2004

C:Accession: T15525

R:Bentley, D.

submitted to the EMBL Data Library, November 1995

A:Description: The sequence of C. elegans cosmid C16B8.

A:Reference number: Z18365

A:Accession: T15525

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-164 <BEN>

A:Cross-references: UNIPROT:Q18043; UNIPARC:UPI000007BFA2; EMBL:U41031; NID:g1098982; PI

C:Genetics:

A:Gene: CESP:C16B8.3

C:Superfamily: Proline-rich peptide P-B

Query Match 31.2%; Score 51.5; DB 2; Length 164;
Best Local Similarity 45.8%; Pred. No. 15;
Matches 11; Conservative 5; Mismatches 7; Indels 1; Gaps 1;

Qy 1 PIGNGNVNGNHLPPAPLPSPQ 24
Db 67 PMG-IGAGVNPSPGYVOQAPPLPTQ 89

RESULT 36

T22685

hypothetical protein F55A11.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C:Accession: T22685

R:Kershaw, J.

submitted to the EMBL Data Library, May 1996

A:Reference number: Z19600

A:Accession: T22685

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Residues: 1-166 <WIL>
A:Cross-references: UNIPROT:Q20796; UNIPARC:UPI0000076858; EMBL:Z72511; PIDN:CAA96655.1
A:Experimental source: clone F55A11
C:Genetics:
A:Gene: CESP:F55A11.1
A:Map position: 5
A:Introns: 33/3; 59/2; 146/2

Query Match 31.2%; Score 51.5; DB 2; Length 166;
Best Local Similarity 50.0%; Pred. No. 17;
Matches 10; Conservative 2; Mismatches 5; Indels 3; Gaps 1;

Qy 7 NNVNGNHLPPAPLPSPQTD 26
Db 120 NNANNNH---QPPPLPSEVE 136

RESULT 37

PRVMM

HIV-1 retropepsin (EC 3.4.23.16) - mouse mammary tumor virus

C:Species: mouse mammary tumor virus, MMTV

C>Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 09-Jul-2004

C:Accession: B26795; A45125

R:Moore, R.; Dixon, M.; Smith, R.; Peters, G.; Dickson, C.

J. Virol. 61, 480-490, 1987

A:Title: Complete nucleotide sequence of a milk-transmitted mouse mammary tumor virus:

A:Reference number: A93030; MUID:87112944; PMID:3027377

A:Accession: B26795

A:Molecule type: DNA

A:Residues: 1-274 <MOO>

A:Cross-references: UNIPROT:PI0271; UNIPARC:UPI0000138D01; EMBL:M15122

R:Menendez-Arias, L.; Young, M.; Oroszlan, S.

J. Biol. Chem. 267, 24134-24139, 1992

A:Title: Purification and characterization of the mouse mammary tumor virus protease ex

A:Reference number: A45125; MUID:93054787; PMID:1331110

A:Accession: A45125

A:Molecule type: DNA

A:Residues: 135-141, 'T', 143-173, 'S', 175-257, 'E', 259-274 <MEN>

A:Cross-references: UNIPARC:UPI000010EC69; GB:I01464; NID:g332105; PIDN:AAA46538.1; PID

A:Note: sequence extracted from NCBI backbone (NCBIN:119080, NCBIPI:119081)

C:Genetics:

A:Gene: prt

C:Complex: homodimer

C:Keywords: aspartic proteinase; homodimer; hydrolase

F:185/Active site: Asp (shared with dimeric partner) #status predicted

Query Match 31.2%; Score 51.5; DB 2; Length 274;
Best Local Similarity 44.8%; Pred. No. 27;
Matches 13; Conservative 3; Mismatches 8; Indels 5; Gaps 1;

Qy 6 GNVNNGNHLPPAP----LPSPQTDGA 29
Db 17 GDGVKSGNLNPEAPPTHTDLPRTGSA 45

RESULT 38

T48825

hypothetical protein 68B2.90 [imported] - Neurospora crassa

C:Species: Neurospora crassa

C>Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 09-Jul-2004

C:Accession: T48825

R:Schulte, U.; Algen, V.; Hohnsbeil, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,

submitted to the Protein Sequence Database, April 2000

A:Reference number: Z24541

A:Accession: T48825

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-503 <SCH>

A:Cross-references: UNIPROT:Q96UB6; UNIPARC:UPI000017B44B; EMBL:AL353821; GSPDB:GN00112,

A:Experimental source: cosmid contig 68B2; strain 74

C:Genetics:

A:Gene: NCSP:68B2.90

A:Map position: 2

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 1, 2006, 22:02:25 ; Search time 199 Seconds
(without alignments)
1282.044 Million cell updates/sec

Title: US-09-189-415D-11
Perfect score: 2840
Sequence: 1 MPIGNIGHNPVNNISIPAP.....SNSAVNTSNPPAPGSHRFV 558

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2589679 seqs, 457216429 residues

Total number of hits satisfying chosen parameters: 2589679

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq.8:*
1: geneseqp19808:*
2: geneseqp19908:*
3: geneseqp20008:*
4: geneseqp20018:*
5: geneseqp20028:*
6: geneseqp20038:*
7: geneseqp20038:*
8: geneseqp20048:*
9: geneseqp20058:*
10: geneseqp20068:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2840	100.0	558	7	ADC00799 Enterohae
2	2840	100.0	558	9	AEB91310 Microbial
3	2840	100.0	558	10	AEE86220 Escherich
4	2624.5	92.4	559	2	AAV06221 EHEC E. c
5	1534.5	54.0	549	2	AAV06220 EPEC E. c
6	406	14.3	107	3	AAAB20576 Intimin C
7	187	6.6	1300	4	ABE68075 Drosophil
8	182.5	6.4	2586	4	ABE68078 Drosophil
9	180	6.3	596	3	AAV99408 Human PRO
10	180	6.3	596	4	AAE66157 Protein o
11	180	6.3	596	4	AAU29178 Human PRO
12	180	6.3	596	4	AAE87575 Human PRO
13	180	6.3	596	4	ABG95900 Human sec
14	180	6.3	596	6	ABU58554 Human PRO
15	180	6.3	596	6	ABU88102 Novel hum
16	180	6.3	596	6	ABU84417 Human sec
17	180	6.3	596	6	ABR66291 Human sec
18	180	6.3	596	6	ABR65681 Human sec
19	180	6.3	596	6	ABU99621 Human sec
20	180	6.3	596	6	ABU82860 Human PRO
21	180	6.3	596	6	ABU89981 Novel hum
22	180	6.3	596	6	ABR68230 Human sec
23	180	6.3	596	6	ABU96283 Novel hum

24	180	6.3	596	6	ABU92714 Human sec
25	180	6.3	596	6	ABO08791 Human sec
26	180	6.3	596	6	ABO02843 Human sec
27	180	6.3	596	6	ABR74997 Human sec
28	180	6.3	596	6	ABR94759 Human sec
29	180	6.3	596	6	ABU85732 Human PRO
30	180	6.3	596	6	ABU98892 Novel hum
31	180	6.3	596	6	ABU98107 Novel hum
32	180	6.3	596	6	ABU91813 Novel hum
33	180	6.3	596	6	ABU89506 Human PRO
34	180	6.3	596	6	ABU86347 Human sec
35	180	6.3	596	6	ABU67560 Human sec
36	180	6.3	596	6	ABU60588 Human PRO
37	180	6.3	596	6	ABU90925 Novel hum
38	180	6.3	596	6	ABO33984 Human sec
39	180	6.3	596	6	ABR99506 Human sec
40	180	6.3	596	6	ABR98896 Human sec
41	180	6.3	596	6	ABO16419 Human sec
42	180	6.3	596	6	ABR92319 Human sec
43	180	6.3	596	6	ABO18960 Human sec
44	180	6.3	596	6	ABR78381 Human sec
45	180	6.3	596	6	ABU72001 Novel hum

ALIGNMENTS

RESULT 1	ADC00799	standard; protein; 558 AA.
ID	ADC00799	
XX	ADC00799;	
XX	04-DEC-2003 (first entry)	
DE	Enterohaemorrhagic E. coli O157:H7-specific protein SEQ ID NO: 844.	
XX	enterohaemorrhagic; anti-bacterial.	
XX	Escherichia coli; O157:H7.	
OS	Escherichia coli; O157:H7.	
XX	JP2002355074-A.	
PN	10-DEC-2002.	
PD	24-JAN-2002; 2002JP-00015959.	
XX	24-JAN-2001; 2001JP-00112010.	
XX	(UTS-) UNIV TSUKUBA.	
PA	WPI; 2003-451640/43.	
XX	Enterohaemorrhagic Escherichia coli O157:H7-specific nucleic acid molecule	
DR	and a polypeptide and its use, a polypeptide, a vector and a host cell.	
XX	Claim 3; SEQ ID NO 844; 2067pp; Japanese.	
PS	The invention relates to a novel enterohaemorrhagic Escherichia coli	
XX	O157:H7-specific nucleic acid molecule. A polynucleotide of the invention	
CC	has anti-bacterial activity. The polypeptide can be used in detection	
CC	and/or treatment of O157:H7 infection. The nucleotide sequence of the	
CC	genome of Enterohaemorrhagic E. coli O157:H7 was determined. The present	
CC	sequence represents an E. coli O157:H7-specific polypeptide of the	
CC	invention.	
XX	Sequence 558 AA;	
SQ	Query Match 100.0%; Score 2840; DB 7; Length 558;	
QY	Best Local Similarity 100.0%; Pred. No. 1,7e-193;	
	Matches 558; Conservativity 0; Mismatches 0; Indels 0; Gaps 0;	
	1 MPIGNIGHNPVNNISIPAPLPUSOTDAGGAGGGLINSTGPGSNALFTPVNNSMADSGD 60	

```

Db      1 MPIGNLGNPNVNNISIPAPLPISQTDGAGRGQLINSTGPLGSRALFTPVNNSMADSGD 60
Qy      61 NRASVPEGLPVNPKRLASETTLNDGFEVLHDHGLDITLNQIGSVFRVETQEDGKHIA 120
Db      61 NRASVPEGLPVNPKRLASETTLNDGFEVLHDHGLDITLNQIGSVFRVETQEDGKHIA 120
Qy      121 VGORNGVTSVYLSDOEYARLOSIDPEGDKRFVFGGKGAGHAMVTVASDITTEARORIL 180
Db      121 VGORNGVTSVYLSDOEYARLOSIDPEGDKRFVFGGKGAGHAMVTVASDITTEARORIL 180
Qy      181 ELLEPKGTGESKGAESKGVGELRESNCGAENTTETQSTSTSLRSDPKMLALGTVAAT 240
Db      181 ELLEPKGTGESKGAESKGVGELRESNCGAENTTETQSTSTSLRSDPKMLALGTVAAT 240
Qy      241 GLIGLAATGIYQALALTRPEPSPPTTDDPAAASATETATRDQLTEAFQNPNOQKVNIDE 300
Db      241 GLIGLAATGIYQALALTRPEPSPPTTDDPAAASATETATRDQLTEAFQNPNOQKVNIDE 300
Qy      301 LGNAIPSGVLKDDVVANIEBOAKAAGEAKQOAIENNAOAKKYDEOQAKROEBLKVSSG 360
Db      301 LGNAIPSGVLKDDVVANIEBOAKAAGEAKQOAIENNAOAKKYDEOQAKROEBLKVSSG 360
Qy      361 AGYGLSGALIIGGGIGVAVTAAALHRKNQPVETQTTTTTTTTTSARTVENKPNANTPAQG 420
Db      361 AGYGLSGALIIGGGIGVAVTAAALHRKNQPVETQTTTTTTTTTSARTVENKPNANTPAQG 420
Qy      421 NVDTGSEDTMESRRSSMASTSTSTFFDTSIGTVQNPYADVKTSLHDSQVPTSNSTSVQ 480
Db      421 NVDTGSEDTMESRRSSMASTSTSTFFDTSIGTVQNPYADVKTSLHDSQVPTSNSTSVQ 480
Qy      481 NMGNNTDSVYVSTIQHPPRDTTNGARLLGNPSAGIQSTYARLALSGGLRHDGGLTGGSN 540
Db      481 NMGNNTDSVYVSTIQHPPRDTTNGARLLGNPSAGIQSTYARLALSGGLRHDGGLTGGSN 540
Qy      541 SAVNTSNPNPAPGSHRFV 558
Db      541 SAVNTSNPNPAPGSHRFV 558

RESULT 2
AEB91310
ID      AEB91310 standard; protein; 558 AA.
AC      AEB91310;
DT      20-OCT-2005 (first entry)
XX      Microbial pathogen adhesin protein sequence, SEQ ID NO:20.
KW      algorithm; adhesin; pharmaceutical; vaccine; drug screening;
KW      bordetella pertussis infection; antibacterial; pneumonia;
KW      antiinflammatory; respiratory-gen.; gastric ulcer; antidiacer;
KW      gastrointestinal-gen.; urinary tract infection; antimicrobial; uropathic.
OS      Escherichia coli.
PN      WO2005076010-A2.
PD      18-AUG-2005.
PF      07-FEB-2005; 2005MO-IN000037.
PR      06-FEB-2004; 2004IN-DE000173.
PR      20-JUL-2004; 2004US-0589227P.
PA      (COUL ) COUNCIL SCI & IND RES SOUTH AFRICA.
PI      Sachdeva G, Kumar K, Jain P, Brahmachari SK, Ramachandran S;
XX      WPI; 2005-597835/61.
XX      Computational method for identifying adhesin and adhesin like molecules,
PT

```

```

PT      comprises computing sequence-based attributes of protein sequences using
PT      neural network software and training an artificial neural network.
PS      Claim 16; SEQ ID NO 20; 402pp; English.
XX      The present invention relates to a computational method (M1) for
XX      identifying adhesin and adhesin-like proteins, by computing the sequence-
XX      based attributes of protein sequences using five attribute modules of a
XX      neural network software, training an artificial neural network (ANN) for
XX      each of the computed five attributes, and identifying the adhesin and
XX      adhesin-like proteins having probability of being an adhesin (Pad) as
XX      equal or greater than 0.51. Also claimed is a set of 274 annotated genes
XX      encoding adhesin and adhesin-like proteins, having 274 fully defined 162-
XX      1151 base pair (SEQ ID NO: 385-658) sequences; a set of 105 hypothetical
XX      genes encoding adhesin and adhesin-like proteins, having 105 fully
XX      defined 306-15876 base pairs (SEQ ID NO: 659-763); a set of 279 annotated
XX      adhesin and adhesin-like proteins, having 279 fully defined 53-3716 base
XX      pair (SEQ ID NO: 1-279); a set of 105 hypothetical adhesin and adhesin-
XX      like proteins, having 105 fully defined 106-5291 base pair (SEQ ID NO:
XX      280-384) sequences; and a fully connected multilayer feed forward ANN (1)
XX      based on (M1). (M1) is useful for identifying adhesin and adhesin-like
XX      proteins, of therapeutic potential, and identifying and short-listing
XX      proteins for further testing in development of new vaccine formulations
XX      to eliminate diseases caused by various pathogenic organisms. (M1) is
XX      useful for identifying putative adhesins that are important in drug
XX      discovery and preventing therapeutics for whooping cough, pneumonia,
XX      gastric ulcer and urinary tract infections. (M1) identifies adhesins from
XX      distantly related organisms, and from bacteria belonging to a wide
XX      phylogenetic spectrum. (M1) is capable of predicting adhesins nature of
XX      unique proteins. The present sequence is a microbial pathogen adhesin
XX      protein sequence.
SQ      Sequence 558 AA;
Qy      Query Match      100.0%; Score 2840; DB 9; Length 558;
Qy      Best Local Similarity 100.0%; Pred. No. 1,76-193;
Qy      Matches 558; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db      1 MPIGNLGNPNVNNISIPAPLPISQTDGAGRGQLINSTGPLGSRALFTPVNNSMADSGD 60
Qy      1 MPIGNLGNPNVNNISIPAPLPISQTDGAGRGQLINSTGPLGSRALFTPVNNSMADSGD 60
Db      61 NRASVPEGLPVNPKRLASETTLNDGFEVLHDHGLDITLNQIGSVFRVETQEDGKHIA 120
Qy      61 NRASVPEGLPVNPKRLASETTLNDGFEVLHDHGLDITLNQIGSVFRVETQEDGKHIA 120
Db      61 NRASVPEGLPVNPKRLASETTLNDGFEVLHDHGLDITLNQIGSVFRVETQEDGKHIA 120
Qy      121 VGORNGVTSVYLSDOEYARLOSIDPEGDKRFVFGGKGAGHAMVTVASDITTEARORIL 180
Db      121 VGORNGVTSVYLSDOEYARLOSIDPEGDKRFVFGGKGAGHAMVTVASDITTEARORIL 180
Qy      121 VGORNGVTSVYLSDOEYARLOSIDPEGDKRFVFGGKGAGHAMVTVASDITTEARORIL 180
Db      121 VGORNGVTSVYLSDOEYARLOSIDPEGDKRFVFGGKGAGHAMVTVASDITTEARORIL 180
Qy      181 ELLEPKGTGESKGAESKGVGELRESNCGAENTTETQSTSTSLRSDPKMLALGTVAAT 240
Db      181 ELLEPKGTGESKGAESKGVGELRESNCGAENTTETQSTSTSLRSDPKMLALGTVAAT 240
Qy      181 ELLEPKGTGESKGAESKGVGELRESNCGAENTTETQSTSTSLRSDPKMLALGTVAAT 240
Db      181 ELLEPKGTGESKGAESKGVGELRESNCGAENTTETQSTSTSLRSDPKMLALGTVAAT 240
Qy      241 GLIGLAATGIYQALALTRPEPSPPTTDDPAAASATETATRDQLTEAFQNPNOQKVNIDE 300
Db      241 GLIGLAATGIYQALALTRPEPSPPTTDDPAAASATETATRDQLTEAFQNPNOQKVNIDE 300
Qy      241 GLIGLAATGIYQALALTRPEPSPPTTDDPAAASATETATRDQLTEAFQNPNOQKVNIDE 300
Db      241 GLIGLAATGIYQALALTRPEPSPPTTDDPAAASATETATRDQLTEAFQNPNOQKVNIDE 300
Qy      301 LGNAIPSGVLKDDVVANIEBOAKAAGEAKQOAIENNAOAKKYDEOQAKROEBLKVSSG 360
Db      301 LGNAIPSGVLKDDVVANIEBOAKAAGEAKQOAIENNAOAKKYDEOQAKROEBLKVSSG 360
Qy      301 LGNAIPSGVLKDDVVANIEBOAKAAGEAKQOAIENNAOAKKYDEOQAKROEBLKVSSG 360
Db      301 LGNAIPSGVLKDDVVANIEBOAKAAGEAKQOAIENNAOAKKYDEOQAKROEBLKVSSG 360
Qy      361 AGYGLSGALIIGGGIGVAVTAAALHRKNQPVETQTTTTTTTTTSARTVENKPNANTPAQG 420
Db      361 AGYGLSGALIIGGGIGVAVTAAALHRKNQPVETQTTTTTTTTTSARTVENKPNANTPAQG 420
Qy      361 AGYGLSGALIIGGGIGVAVTAAALHRKNQPVETQTTTTTTTTTSARTVENKPNANTPAQG 420
Db      361 AGYGLSGALIIGGGIGVAVTAAALHRKNQPVETQTTTTTTTTTSARTVENKPNANTPAQG 420
Qy      421 NVDTGSEDTMESRRSSMASTSTSTFFDTSIGTVQNPYADVKTSLHDSQVPTSNSTSVQ 480
Db      421 NVDTGSEDTMESRRSSMASTSTSTFFDTSIGTVQNPYADVKTSLHDSQVPTSNSTSVQ 480
Qy      421 NVDTGSEDTMESRRSSMASTSTSTFFDTSIGTVQNPYADVKTSLHDSQVPTSNSTSVQ 480
Db      421 NVDTGSEDTMESRRSSMASTSTSTFFDTSIGTVQNPYADVKTSLHDSQVPTSNSTSVQ 480
Qy      481 NMGNNTDSVYVSTIQHPPRDTTNGARLLGNPSAGIQSTYARLALSGGLRHDGGLTGGSN 540
Db      481 NMGNNTDSVYVSTIQHPPRDTTNGARLLGNPSAGIQSTYARLALSGGLRHDGGLTGGSN 540

```

QY 541 SAVNTSNNPPAPGSHRFV 558
|||
DB 541 SAVNTSNNPPAPGSHRFV 558

RESULT 3
AEE86220
ID AEE86220 strand; protein; 558 AA.
XX
AC AEE86220;
XX
DT 23-FEB-2006 (first entry)
XX
DE Escherichia coli translocated intimin receptor (Tir) protein.
XX
KW Diagnosis; therapeutic; screening; escherichia coli infection;
KW antibacterial; infection; translocated intimin receptor.
XX
OS Escherichia coli.
XX
PN US2005287569-A1.
XX
PD 29-DEC-2005.
XX
PF 20-MAY-2005; 2005US-00134563.
XX
PR 20-MAY-2004; 2004US-0573600P.
XX
PA (LEON/) LEONG J M.
PA (CAMP/) CAMPBELLONE K G.
XX
PI Leong JM, Campellone KG;
XX
DR WPI; 2006-065745/07.
DR N-PSDB; AEE86219.
XX
PT Novel purified polypeptide having six residues of EspF-U, and binding to
PT neuronal Wiskott-Aldrich syndrome protein polypeptide, useful for
PT identifying EspF-U activity modulating compound.
XX
PS Disclosure; SEQ ID NO 12; 62pp; English.
XX
XX
CC The present invention relates to novel EspFU polypeptides and their
CC corresponding polynucleotides. The EspFU polypeptides are EspF-like
CC polypeptides encoded by genes of the cryptic prophage CP-933U of
CC enterohemorrhagic Escherichia coli (EHEC) or enteropathogenic Escherichia
CC coli (EPEC) that binds to a neuronal wiskott-aldrich syndrome protein (N-
CC WASP) polypeptide or restores the actin pedestal formation activity of
CC enteropathogenic E. coli (EPEC) strain KC12. The invention further
CC relates to a method of identifying a candidate compounds capable of
CC binding to and/or modulating the activity of EspFU and compounds that
CC inhibit protein-protein interactions between EspFU and EspFU-interacting
CC proteins such as N-WASP, transducer of Cdc42-dependent actin assembly-1
CC (Toca-1) and p21-activated kinase 1 (pak1). EspFU polynucleotides are
CC useful for diagnosing or detecting EHEC infection. EspFU antibody is
CC useful for treating EHEC infection. The present sequence is the
CC enterohemorrhagic Escherichia coli translocated intimin receptor (Tir)
CC protein. This sequence is critical for the formation of actin pedestals
CC in EHEC.
XX
SQ Sequence 558 AA;

Query Match 100.0%; Score 2840; DB 10; Length 558;
Best Local Similarity 100.0%; Pred. No. 1.7e-193;
Matches 558; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRIGNGHNPNNSGTPAPPLPSQTDGAGRGQLINSGPLGSRALLFPVNSMADSGD 60
|||
DB 1 MRIGNGHNPNNSGTPAPPLPSQTDGAGRGQLINSGPLGSRALLFPVNSMADSGD 60
|||

QY 61 NRASDVPGLPVNPMLAASEITLNDGFVYLHDHGPLDTLNROIGSSVFVETQEDGKHIA 120
|||

DB 61 NRASDVPGLPVNPMLAASEITLNDGFVYLHDHGPLDTLNROIGSSVFVETQEDGKHIA 120
|||
QY 121 VQQRNVEVTSVVLSDQFYARLOSIDEQKDKFTFGRGGAGAMTTVASDITEARQIL 180
|||
DB 121 VQQRNVEVTSVVLSDQFYARLOSIDEQKDKFTFGRGGAGAMTTVASDITEARQIL 180
|||
QY 181 ELLEPKGTESKAGSKGVGELRESNSGAENTTETQTSSTSLRSDPKLMLALGVAT 240
|||
DB 181 ELLEPKGTESKAGSKGVGELRESNSGAENTTETQTSSTSLRSDPKLMLALGVAT 240
|||
QY 241 GLIGLAATGIVQALALTPPEPDSPTTDPDAASATETATRDQLTKEAFQPNQKNVIDE 300
|||
DB 241 GLIGLAATGIVQALALTPPEPDSPTTDPDAASATETATRDQLTKEAFQPNQKNVIDE 300
|||
QY 301 LGNAIPSGVLKDDVYANIEQAKAGEAKQQAENNAQOKKYDEQAKROELKVSIG 360
|||
DB 301 LGNAIPSGVLKDDVYANIEQAKAGEAKQQAENNAQOKKYDEQAKROELKVSIG 360
|||
QY 361 AGYGLSGALILGGGIGVAVTALHKNQPVQETTTTTTTTTTARTVENKPPANNTPAOG 420
|||
DB 361 AGYGLSGALILGGGIGVAVTALHKNQPVQETTTTTTTTTTARTVENKPPANNTPAOG 420
|||
QY 421 NVDTPESEDPTMESRRSSMASTSTFPDTSISIGVQNPVADVKTSLHDSQVPTSNSTSVQ 480
|||
DB 421 NVDTPESEDPTMESRRSSMASTSTFPDTSISIGVQNPVADVKTSLHDSQVPTSNSTSVQ 480
|||
QY 481 NMGNTDSVVYSTIQHPPRDITDNGARLLGNPSAGISTYARIALSGGLRHDMGGLTGGSN 540
|||
DB 481 NMGNTDSVVYSTIQHPPRDITDNGARLLGNPSAGISTYARIALSGGLRHDMGGLTGGSN 540
|||

QY 541 SAVNTSNNPPAPGSHRFV 558
|||
DB 541 SAVNTSNNPPAPGSHRFV 558

RESULT 4
AA06221
ID AA06221 strand; protein; 559 AA.
XX
AC AA06221;
XX
DT 16-AUG-1999 (first entry)
XX
XX
DE EHEC E. coli translocated intimin receptor (Tir).
XX
KW Tir; translocated intimin receptor; H90; enterohaemorrhagic; EHEC;
KW infection; diagnosis; vaccine.
XX
OS Escherichia coli.
XX
XX
FH Key Location/Qualifiers
FT Misc-difference 453
FT /note= "encoded by codon of 1 apparent nucleotide,
FT causing frameshift in the DNA sequence"

PN WO9924576-A1.
XX
XX 20-MAY-1999.
XX
PD 10-NOV-1998; 98WO-CA001042.
XX
PF 12-NOV-1997; 97US-0065130P.
XX
PR (UYBR-) UNIV BRITISH COLUMBIA.
XX
PA Finlay BB, Kenny B, Deviney R, Stein M;
PI WPI; 1999-337712/28.
XX
DR N-PSDB; AAX58859.
XX
PT New translocated intimin receptor useful for treating infection by
PT enteropathogenic or enterohemorrhagic Escherichia coli.
XX

PS Claim 7, Page 55-58; 91pp; English.
XX The present sequence represents Tir, a novel translocated intimin
CC receptor (formerly termed Hp90) from an enterohaemorrhagic Escherichia
CC coli (EHEC) strain. The sequence was deduced from an isolated tir
CC polynucleotide (see AAX58859). Tir proteins are secreted by attaching and
CC effecting pathogens such as EHEC and EPEC (see AAY06220) E. coli. The
CC bacterial pathogens insert their own receptors into mammalian cell
CC surfaces, to which the pathogen then adheres to trigger additional host
CC signaling events and actin nucleation. Diagnosis of disease caused by
CC pathogenic E. coli can be performed by use of antibodies that bind to Tir
CC to detect the protein or the use of nucleic acid probes for detection of
CC nucleic acids encoding Tir polypeptide. Isolated Tir nucleic acids, Tir
CC peptides, a recombinant method for producing recombinant Tir, antibodies
CC which bind to Tir, and a kit for the detection of Tir-producing E. coli
CC are provided. A method of immunising a host with Tir to induce a
CC protective immune response is also provided. In addition, Tir fusion
CC proteins can be used in attenuated E. coli to induce a cell-mediated
CC immune response to other polypeptides, e.g. antigens. A method for
CC screening for compounds which interfere with the binding of bacterial
CC pathogens to their receptors is further provided
XX

SQ Sequence 559 AA;

Query Match 92.4%; Score 2624.5; DB 2; Length 559;
Best Local Similarity 94.0%; Pred. No. 3,8e-178;
Matches 528; Conservative 3; Mismatches 24; Indels 7; Gaps 3;

QY 1 MPIGLGNPNVNSIPPAPLPSQTDGAGRGQLINSTGPIGRALFTPVNSMADSGD 60
DB 1 MPIGLGNPNVNSIPPAPLPSQTDGAGRGQLINSTGPIGRALFTPVNSMADSGD 60
QY 61 NRASVPGI.PVNPMLAA--SEITLNDGFEVLHDHGPLDTLNRIQSSVFRVETQEDGKHA 120
DB 61 NRASVPGI.PVNPMLAA--SEITLNDGFEVLHDHGPLDTLNRIQSSVFRVETQEDGKHA 120
QY 121 VGORNGVETSVLSDOEVARLQSIDPEGKDFVTGKSGAGAAHAAVTAASDITEARQIL 180
DB 121 VGORNGVETSVLSDOEVARLQSIDPEGKDFVTGKSGAGAAHAAVTAASDITEARQIL 180
QY 181 ELLEBKGTGESKGAESKGVGELRSNSGAENTTETQISTSTSLRSDPKMLALGVAT 240
DB 181 ELLEBKGTGESKGAESKGVGELRSNSGAENTTETQISTSTSLRSDPKMLALGVAT 240
QY 241 GLIGLAAGTIVQALALTPEPDSPTTTPDPAASATETARLDLTKEAPNPNOKVNI 300
DB 241 GLIGLAAGTIVQALALTPEPDSPTTTPDPAASATETARLDLTKEAPNPNOKVNI 300
QY 301 LGNAIPSGVLKDDVAVANIEEQAAGAEAKQAENNAQAQKYDEQAQKROEELKVS 360
DB 301 LGNAIPSGVLKDDVAVANIEEQAAGAEAKQAENNAQAQKYDEQAQKROEELKVS 360
QY 361 AGYGLSGALLIGGIGVAVTAALHKNQPEQTTTTTTTTTSARVYENKPNANTPAQ 420
DB 361 AGYGLSGALLIGGIGVAVTAALHKNQPEQTTTTTTTTTSARVYENKPNANTPAQ 420
QY 421 NVDTGSGEDTMSRSSMASTSTFPPDSSIGTVONPVADVTSI--HDSQVTSNSISV 479
DB 421 NVDTGSGEDTMSRSSMASTSTFPPDSSIGTVONPVADVTSI--HDSQVTSNSISV 479
QY 480 QNM---GNTDSVYVYETIQHPPTDINGARLIGNPSAGIOSTYARLALSGLRHMGGLT 536
DB 480 QNM---GNTDSVYVYETIQHPPTDINGARLIGNPSAGIOSTYARLALSGLRHMGGLT 536
QY 478 RLFRITWGQISVYVETIQHPPTDINGARLIGNPSAGIOSTYARLALSGLRHMGGLT 537
DB 478 RLFRITWGQISVYVETIQHPPTDINGARLIGNPSAGIOSTYARLALSGLRHMGGLT 537
QY 537 GGSNSAVNTSNPPAPGSHRFV 558
DB 537 GGSNSAVNTSNPPAPGSHRFV 558
QY 538 GGSNSAVNTSNPPAPGSHRFV 559
DB 538 GGSNSAVNTSNPPAPGSHRFV 559

RESULT 5
AAY06220
ID AAY06220 standard; protein; 549 AA.
XX

AC AAY06220;
XX
XX 16-AUG-1999 (first entry)
XX
XX EPEC E. coli translocated intimin receptor (Tir).
DE
XX Tir; translocated intimin receptor; Hp90; enteropathogenic; EPEC;
KW infection; diagnosis; vaccine.
XX
XX Escherichia coli.
OS
XX
XX
FH Key Location/Qualifiers
XX
XX Misc-difference 180
FT /note="encoded by AAA"
FT
FT Domain.
FT 234..253
FT /note="putative transmembrane domain"
FT
FT Misc-difference 314
FT /note="given as Xaa in the specification; Lys is deduced
FT from the DNA sequence"
FT 364..386
FT Domain
FT /note="putative transmembrane domain"

XX WO9924576-A1.
XX
XX 20-MAY-1999.
XX
XX 10-NOV-1998; 98WO-CA001042.
XX
XX 12-NOV-1997; 97US-0065130P.
XX
XX (UYBR-) UNIV BRITISH COLUMBIA.
XX
XX Finlay BB, Kenny B, Devlin R, Stein M;
XX WPI; 1999-337712/28.
XX
XX N-PSDB; AAX58858.
XX
XX New translocated intimin receptor useful for treating infection by
XX enteropathogenic or enterohaemorrhagic Escherichia coli.
XX
XX
PS Claim 6, Page 55-58; 91pp; English.
XX
XX The present sequence represents Tir, a novel translocated intimin
CC receptor (formerly termed Hp90) from an enteropathogenic Escherichia coli
CC (EPEC) strain. The sequence was deduced from an isolated tir
CC polynucleotide (see AAX58858). Tir proteins are secreted by attaching and
CC effecting pathogens such as EPEC and EHEC (see AAY06221) E. coli. The
CC bacterial pathogens insert their own receptors into mammalian cell
CC surfaces, to which the pathogen then adheres to trigger additional host
CC signaling events and actin nucleation. Diagnosis of disease caused by
CC pathogenic E. coli can be performed by use of antibodies that bind to Tir
CC to detect the protein or the use of nucleic acid probes for detection of
CC nucleic acids encoding Tir polypeptide. Isolated Tir nucleic acids, Tir
CC peptides, a recombinant method for producing recombinant Tir, antibodies
CC which bind to Tir, and a kit for the detection of Tir-producing E. coli
CC are provided. A method of immunising a host with Tir to induce a
CC protective immune response is also provided. In addition, Tir fusion
CC proteins can be used in attenuated E. coli to induce a cell-mediated
CC immune response to other polypeptides, e.g. antigens. A method for
CC screening for compounds which interfere with the binding of bacterial
CC pathogens to their receptors is further provided
XX

SQ Sequence 549 AA;

Query Match 54.0%; Score 1534.5; DB 2; Length 549;
Best Local Similarity 56.4%; Pred. No. 1.6e-100;
Matches 328; Conservative 60; Mismatches 137; Indels 57; Gaps 11;

QY 1 MPIGLGNPNVNSIPPAPLPSQTDGAGRGQLINSTGPIGRALFTPVNSMADSGD 58
DB 1 MPIGLGNPNVNSIPPAPLPSQTDGAGRGQLINSTGPIGRALFTPVNSMADSGD 60
QY 59 GDNRASDVPGI.PVNPMLAA--SEITLNDGFEVLHDHGPLDTLNRIQSSVFRVETQEDG 116

```

Db      61  VDSR--DIPGLPTNPRLAATSETCLLGGFEVLHDKGPDILINTQIGPSAFVEVQADG 118
QY      117 KHLAVQQRNVERSVLVSDDEVARLQSIDPEGKDKPFYTCGRGAGAMTVASDITEAR 176
Db      119 THAIEKNGLEVSVTLSPQEWSSLOSIDEGKRRFVFYTCGRGSGHPMTTVASDIAEAR 178
QY      177 ORILELEPKGTG-----ESKAGESEKGVGELRESNGAENETOTSTSTLSRS 227
Db      179 TRILATLDPNNGRGKPKVDVTRSVGASAGI-----DDGV--VSEHTSTSTSSVRS 230
QY      228 DPKLWALGTVAATGLIGLAATGIVQALALTPBEDSPPTTDPDAASATETATRDOLTKEA 287
Db      231 DPKFVWSVGAIAAGLAGLAATGIAQALALTPBEDDPTTDPDAANAESATDQLOLQEA 290
QY      288 FQNPNDQKNIDELGNALPSGVLDKDVVANIEBOAKAAGEAQOAIENNAQOKKYDEQ 347
Db      291 FKNPENQKNIDANGNALPSGELKDDIVEQIAQOAKEAGEVANAQAVESNAQOKRYEDQ 350
QY      348 QAKROBELKVSAGVGLSGALILGGIGVAVTAALHRRKQPVQOTTPTTTTTTTSART 407
Db      351 HARROBELQSSGIGLSSALIVAGSIGAVTTALHRRKQPAEQTTT-----HT 403
QY      408 VENKPNANTPAQGNVDTPGSEDTMESRRSSMASTSTSPFPTSIGTVQNPYADV-----X 462
Db      404 VVQOQTGSIPOHKVVALMPQERRRFSRDRDSQGSVAATHWSDS--SEVANPYAEVGGARNS 462
QY      463 TSLHD-----SOVPTSNSNTSVQNMGNITDSVYSTIQHPRDTTNGARLLGNPSAGIQ 516
Db      463 LEAHQPEEHIYDEVAADPGSVYIQNFGSGSPVT-----GRLLIGTPGQGIQ 507
QY      517 STYARLAUSGGLRHDWGLTGGSNSAVMNSNPAPGSRFV 558
Db      508 STYALLANSGLRLGMGLTSGGTAVSSVNAAPTPGPVRFV 549

```

```

RESULT 6
AAB20576
ID AAB20576 standard; protein; 107 AA.
XX
AC AAB20576;
XX
DT 08-DEC-2000 (first entry)
XX
DE Intimin C-terminal Tir binding domain amino acid sequence.
XX
KW Intimin; Tir binding domain; inhibitor; intimin adhesion; screening;
KM Tir-independent eukaryotic cell binding activity; bacterial infection;
KW diarrhoea; antibacterial.
XX
OS Unidentified.
XX
PN WO200045173-A1.
XX
PD 03-AUG-2000.
XX
PF 31-JAN-2000; 2000WO-GB000254.
XX
PR 29-JAN-1999; 99GB-00001897.
XX
PA (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.
XX
PI Frankel GM, Matthews SJ, Hale CB, Dougan G;
XX
DR WPI; 2000-499357/44.
XX
PT Screening for inhibitors of intimin binding to eukaryotic cells, for use
PT in diagnosing, preventing and treating bacterial infections, especially
PT Escherichia coli O157:H7.
XX
PS Claim 8; Page 76; 96pp; English.
XX
CC The present invention describes a method of screening for an inhibitor of

```

```

CC      CC      Intimin binding to eukaryotic cells. The method comprises exposing an
CC      CC      intimin polypeptide having a Tir-independent cell binding activity to
CC      CC      test agents, and obtaining an inhibitor based on its ability to bind the
CC      CC      polypeptide. The inhibitors are used in the prevention, treatment and/or
CC      CC      diagnosis of bacterial infections, preferably by enteropathic and/or
CC      CC      enterohaemorrhagic Escherichia coli, Shiga toxinigenic E. coli, Hafnia
CC      CC      alvei or Citrobacter freundii, or especially E. coli O157:H7. The
CC      CC      infections cause a histopathological effect known as attachment and
CC      CC      effacement on intestinal epithelial cells. The inhibitors can be used to
CC      CC      produce food supplements or additives, especially where the food is a
CC      CC      milk substitute. The method can be used to sort cells based on their
CC      CC      ability to bind to a Tir independent cell binding domain of an intimin
CC      CC      polypeptide. Polypeptides having Tir-independent intimin binding activity
CC      CC      can be used to produce a vaccine against a bacterial disease. The present
CC      CC      sequence represents a specifically claimed intimin C-terminal Tir binding
CC      CC      domain amino acid sequence, for use in the method of the present
CC      CC      invention
XX
SQ      Sequence 107 AA;

```

```

Query Match      14.3%; Score 406; DB 3; Length 107;
Best Local Similarity 72.0%; Pred. No. 4,4e-21;
Matches 77; Conservative 18; Mismatches 12; Indels 0; Gaps 0;

```

```

QY      252 QALALTPBEDSPPTTDPDAASATETATRDOLTKFAFQNPNDQKNIDELGNALPSGVTK 311
Db      1 QALALTPBEDDPTTDPDAANAESATDQLOLQEFKNPENQKNIDANGNALPSGELK 60

```

```

QY      312 DVVANIEBOAKAAGEAQOAIENNAQOKRYDEQOAKROBELKYS 358
Db      61 DIVEQIAQOAKEAGEVANAQAVESNAQOKRYEDQHARROBELQS 107

```

```

RESULT 7
AAB68075
ID AAB68075 standard; protein; 1300 AA.
XX
AC AAB68075;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 31017.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KM pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US009231.
XX
PR 23-MAR-2000; 2000US-0191637P.
XX
PR 11-JUL-2000; 2000US-00614150.
XX
PA (PEKE ) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
XX
DR N-PSDB; ABL12178.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX
PS Disclosure; SEQ ID NO 31017; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and

```

cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL130511), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-ABBS72072). The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

Sequence 1300 AA;

Query Match 6.6%; Score 187; DB 4; Length 1300;
Best Local Similarity 20.4%; Pred. No. 0.0047;
Matches 130; Conservative 74; Mismatches 245; Indels 188; Gaps 25;

```
QY 5 NLGNPNVNNNSIPRPLPSQDQ---AGRGQLNSTGRLPLFTVPRMSMDSQDN 61
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 438 NSTSNSNTNDSTGSETSTSTGLVASGAG-----GATGAAMLPTP---SQOSTGAK 488
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 62 RAS-----DVPGLPVNPMRLA-----ASEITLNDGFEVLHDHGLDPLTLNRQ 102
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 489 EATAVSLLEKLPVNVVSPILMKELRQKMTKYDAEMIMANAAVQOQH----- 539
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 103 IGSSVFRVETDQGHIAVGQRNVEVTSVLSDEYARLQSIDPEGKQFVFTGREGAG 162
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 540 --QHFFHHHHHHHHHHHGHASTGAETAAVQOQMAAMQKPG-----VGGTGAAG 588
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 163 HAMTVASDITFARQRIELLEPKGTGSKGSGKGCYGLRESNGAENTETOTST 222
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 569 NAGATTVSSVA-----AAGSEVNGRSTSLRKSRRVNS-----TSSSI 627
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 223 SLSRSDPKLMLALGTAVATGLGLAATGIVQALALTPEDPSPTTDPDAASATETATDQ 282
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 628 STASADVI-----APVVASISLPSKAPVYVLMRCKPQAMATAALHQ 670
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 283 LITKAFQNPNDQKVIDELGNALPSGVLDKDVVNIIEQAKAAGEBAKQ-----A 333
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 671 SQORQLRSEKKEKLTLD-----GESSD--TSSEQOK--EOKQDHLQPKMFS 716
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 334 IENNAQKRYDEQOAKQOELKVSSGAGYGLSGLLILGGIGVAVTALHHRKNQPVQGT 393
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 717 LAEPQPEKS--BEKQEOQKRVTRNSAGRVG-----VARLTAHNN--IA 760
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 394 TTTTNTTTSARVY---ENKEPANTPAQGNVDTGSEDMESRRSSMASTSTFFDTS 450
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 761 TTTNSSSSSKTKTTTTCNNHNSNSRNNHNSLSRLSVSKRPAPSEASSIPSS7SS 820
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 451 IGTVO-----NPYADVTSILHDSQVPTSNSTSVQNMGNTDSVVYSTIOH----- 495
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 821 ENOQOQATRRSCSPYAKKNLASFDPDPPTGCIKEQLKDESVTVSPYKORRRRAA 880
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 496 -----PRDDTTDNGARL--LGNPSAIGST-----YARLALSGG 527
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 881 LAAAGSIHCEALIGFPTSGTSGORRPAQAGEPTTSCSTTISNVEPLKTPERRLKLTLR 940
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 528 LRH-----DMGILTGSGNSAVNTSNPNPAPGSHR 556
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 941 MKRSPILDEVIELG--TSLNCGAGRG---APGSHR 971
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
```

RESULT 8
ABBS6878
ID ABBS6878 standard; protein; 2586 AA.

AC ABBS6878;

XX 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 27426.

KM Drosophila; developmental biology; cell signalling; insecticide;
pharmaceutical.

XX Drosophila melanogaster.

XX MO200171042-A2.
XX 27-SEP-2001.
XX 23-MAR-2001; 2001WO-US009221.
XX 23-MAR-2000; 2000US-0191637P.
XX 11-JUL-2000; 2000US-00614150.
XX (PEKE) PE CORP NY.
XX Venter JC, Adams M, Li PWD, Myers EW;
XX WPI; 2001-656860/75.
XX N-PSDB; ABL10981.

PT New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions.

PS Disclosure; SEQ ID NO 27426; 21pp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL130511), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-ABBS72072). The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 2586 AA;

Query Match 6.4%; Score 182.5; DB 4; Length 2586;
Best Local Similarity 19.4%; Pred. No. 0.0025;
Matches 113; Conservative 85; Mismatches 237; Indels 147; Gaps 19;

```
QY 4 GNLCNPNVNNNSIPRPLP-----SQTDAGGKGLQIN-----STGPLGSRALFTVYR 52
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 71 GELCNPNVYVNCVNGKKQLPLDSSITDHSASTNTTKRIIDGLSSQTTSSAPVVDVT 130
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 53 NSMADSGNRRASDVPGLPVNPMRLAASEITLNDGFEVLHDHGLDPLTLRQIGSSVFRVET 112
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 131 QGSSNGDGNSTQ-----STTTTITTTTSSDGEFTTSSDPVEVSO--GTMGNSST 182
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 113 QEDGKHIAVGQRNVEVTSVLSDEYARLQSIDP-----EKKDFVFTGREGAGHAMVT 167
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 183 QSSS-----STTTTSSDEGQTTSSDPVEVAQSS-----SNGDGNSTQSLT 227
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 168 VASDITFARQRIELLEPKGTGSKGSGKGCYGLRESNGAENTETOTSTST----- 222
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 228 TTTTNTTTS-----SDGGQSTTSSDPVEVSGTNGMSSTQSSSTTTTSSD 275
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 223 ---SLSRSDPKLMLALGTAVATGLGLAATGIVQALALTPEDPSPTTDPDAASATETAT 279
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 276 EGQTTSSDPVEVAQSSNG-----DGNSTQSSSTTTT 310
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 280 RDOLITKAFQNPNDQKVIDELGNALPSGVLDKDVVNIIEQAKAAGEBAKQAIENNAQ 339
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 311 TTTTSSDGGQ-----STTLSDPVV-----EVSQGTNGNNS 341
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 340 AQKTYDEQOAKQOELKVSSGA--GYLSGALLGGIGVAVTALHHRKNQPVHQTTT 396
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 342 TQSSSTTTTSSDEGQTTSSDPVEVAQSSSNGDGNST-----QSSIT 387
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 397 TTTTNTTTSARVYENKPNANTPAQGNVDTGSEDMESRRSSMASTSTFFPTSSIGTYQN 456
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 388 TTTTNTTSSDGGQSTTSSDPVEVSGTNGAN--STQSSATTTTSSDGGQTTSSSD 444
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 457 PYADV---KTLHDSQVPTSNSTSVQNMGNTDSVVYSTIOHPRDDTT--NGARLLGNP 511
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
```



```
Db 108 SSSSTTSSGASTATNBE---SSTPSSGASTVTNCGSSVTSSGASTATNSESSTVSSRAS 163
QY 181 ELLEPKGTGSKAGSGKGVGELRESNGAENTTETOTSTSTSLRSDPLMLALGTVA 240
Db 164 TATNSESSTLSSGASTATN-SDSSTTSSGASTATNSESSTTSSG-----ASTAT 211
QY 241 GLIGLATGIVQALALTPEDPSPTTDPDAASATETATPDQLTKEAFQNDQKVNIDE 300
Db 212 N---SSSTVSSRASATATNSESSTT---SSGASTATNNSRRTTNSGATATNSESSTTS 264
QY 301 LG-----NAIPGCVLKDVANIEEQAKAAGEBAKQOAIENNAQAKKYDEQOAKR 351
Db 265 SGASTATNDSSTVSGA---STATNSESSTTSSGAST-----ATN 302
QY 352 QEELKXSSGAGVYL-SGALLLGGIGVAVTAALHRKNQVPEQTTTTTTTTTSARIVEN 410
Db 303 SSSSTTSSGASTATNDSSTTSSGAGATNSESSTVSSGISTVTNSESSTPSSGANTATN 362
QY 411 KPAANTPAQGNVDPSPSEDPTMESRRSMASTSTFPEDTSIGTVQNPYADVKT-----464
Db 363 SSSSTTSSGANTAT-NSESSTVSSGASTATNSESSTTSSGVTATNSESSTTSSGASTAT 421
QY 465 LHDQVPTNSNTSVONMGNTDSVYVSTIQHPRDITDNGARLLGNPSAGIQSTVARLAL 524
Db 422 NSDSSTTSSSEASATATNSESSTVSSGISTVTNSESSTTSSGANTATNCGSSVTS-----474
QY 525 SGGLRHDMGGLTGGSNSAVNTSNPPAPG 553
Db 475 AGSGTALTGMTTTHSA-STAVSEAKPG 502

RESULT 11
AAU29178
ID AAU29178 standard; protein; 596 AA.
XX
AC AAU29178;
XX
DT 18-DEC-2001 (first entry)
XX
DE Human PRO polypeptide sequence #155.
XX
KW PRO polypeptide; mammal; tumour; cancer; human; cattle; horse; sheep;
KW dog; cat; pig; goat; rabbit; tumour necrosis factor alpha; TNF-alpha;
KW blood; chondrocyte cell; cell proliferation; cell differentiation; colon;
KW adrenal; lung; breast; prostate; rectum; cervix; liver; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200168848-A2.
PD 20-SEP-2001.
XX
XX 28-FEB-2001; 2001WO-US006520.
XX
PR 01-MAR-2000; 2000WO-US005601.
PR 02-MAR-2000; 2000WO-US005841.
PR 03-MAR-2000; 2000US-0187202P.
PR 06-MAR-2000; 2000US-0186968P.
PR 14-MAR-2000; 2000US-0189320P.
PR 14-MAR-2000; 2000US-0189328P.
PR 15-MAR-2000; 2000WO-US006884.
PR 21-MAR-2000; 2000US-0190828P.
PR 21-MAR-2000; 2000US-0191007P.
PR 21-MAR-2000; 2000US-0191048P.
PR 21-MAR-2000; 2000US-0191314P.
PR 28-MAR-2000; 2000US-0192655P.
PR 29-MAR-2000; 2000US-0193032P.
PR 29-MAR-2000; 2000US-0193053P.
PR 30-MAR-2000; 2000WO-US008439.
PR 04-APR-2000; 2000US-0194449P.
PR 04-APR-2000; 2000US-019447P.
PR 11-APR-2000; 2000US-0195975P.
PR 11-APR-2000; 2000US-0196006P.
```

```
PR 11-APR-2000; 2000US-0196187P.
PR 11-APR-2000; 2000US-0196690P.
PR 11-APR-2000; 2000US-01966820P.
PR 18-APR-2000; 2000US-0198121P.
PR 18-APR-2000; 2000US-0198585P.
PR 25-APR-2000; 2000US-0199397P.
PR 25-APR-2000; 2000US-0199550P.
PR 25-APR-2000; 2000US-0199654P.
PR 03-MAY-2000; 2000US-0201516P.
PR 17-MAY-2000; 2000WO-US013705.
PR 22-MAY-2000; 2000WO-US014042.
PR 30-MAY-2000; 2000WO-US014941.
PR 02-JUN-2000; 2000WO-US015264.
PR 05-JUN-2000; 2000US-0209832P.
PR 28-JUL-2000; 2000WO-US020710.
PR 22-AUG-2000; 2000US-00644848.
PR 24-AUG-2000; 2000WO-US023328.
PR 08-NOV-2000; 2000WO-US030952.
PR 01-DEC-2000; 2000WO-US032678.
PR 20-DEC-2000; 2000WO-US034956.

PA (GETH ) GENENTECH INC.
XX
XX Baker KP, Chen J, Desnoyers L, Goddard A, Godoweki PJ, Gurney AL,
PI Pan J, Smith V, Watanabe CK, Wood WI, Zhang Z,
XX
XX N-PSDB; AAS46079.
DR
XX
XX MPI; 2001-602746/68.
PT
PT Novel nucleic acids encoding PRO polypeptides, used to diagnose the
PT presence of tumors, such as prostate and breast tumors, in mammals and to
PT screen for modulators of the compounds.
XX
XX Claim 11; Fig 310; 774pp; English.
XX
XX Sequences AAU29024-AAU29328 represent PRO polypeptides of the invention.
CC The PRO polypeptides and their associated nucleic acids can be used to
CC detect the presence of a tumour in a mammal by comparing the level of
CC expression of a PRO polypeptide in a test sample of cells from the animal
CC and a control sample of normal cells, whereby a higher level of
CC expression in the test sample indicates the presence of a tumour in the
CC mammal. Mammals include dogs, cats, cattle, horses, sheep, pigs, goats
CC and rabbits but are preferably human. The polypeptides can be used to
CC stimulate tumour necrosis factor (TNF) alpha release from human blood,
CC when contacted with it. A specific polypeptide can be used to stimulate
CC the proliferation or differentiation of chondrocyte cells. The PRO
CC proteins can be used to determine the presence of tumours and also
CC susceptibility to tumour development, particularly adrenal, lung, colon,
CC breast, prostate, rectal, cervical, or liver tumours, in mammalian
CC subjects. The oligonucleotide probes specific for the PRO nucleic acids
CC can be used for genetic analysis of individuals with genetic disorders
CC
XX
XX SQ Sequence 596 AA;
XX
XX Query Match 6.3%; Score 180; DB 4; Length 596;
XX Best Local Similarity 19.2%; Pred. No. 0.00053;
XX Matches 109; Conservative 90; Mismatches 272; Indels 98; Gaps 19;
QY 4 GNIGHNPVNNNSIPPAIPLPISQTDGAGRGGLINSTGPGSRALFPVNSNMDSGDNRA 63
Db 13 GLLHLEAATNS-----NETSTSNANTGSSVTSG-----ASTATNG 49
QY 64 SDVPGLPVPMRLAASEITLNDGFEVLHDGCPDLTLNRQIGSSV-PRVETQEDGKHIAVG 122
Db 50 SSVTSSGVTATISGSSVTSN-GVSIV-TNSEPHITSGISTATNSEFSTASGISIATN 107
QY 123 QKNGVETS--VLSDQEVARLQSIDPEGDKRVFTGREGGAGHAMVTASDITAEARQL 180
Db 108 SSSSTTSSGASTATNBE---SSTPSSGASTVTNCGSSVTSSGASTATNSESSTVSSRAS 163
QY 181 ELLEPKGTGSKAGSGKGVGELRESNGAENTTETOTSTSTSLRSDPLMLALGTVA 240
Db 164 TATNSESSTLSSGASTATN-SDSSTTSSGASTATNSESSTTSSG-----ASTAT 211
```


OS	XX	Hom sapiens.	
PN	XX	US2002119130-A1.	
PD	XX	29-AUG-2002.	
PF	XX	06-DEC-2001; 2001US-00006667	
PR	XX	29-OCT-1997;	97US-0063435P
PR	XX	29-OCT-1997;	97US-0064215P
PR	XX	22-APR-1998;	98US-0082797P
PR	XX	29-APR-1998;	98US-0083495P
PR	XX	15-MAY-1998;	98US-0085579P
PR	XX	02-JUN-1998;	98US-0087759P
PR	XX	04-JUN-1998;	98US-0088021P
PR	XX	04-JUN-1998;	98US-0088029P
PR	XX	10-JUN-1998;	98US-0088030P
PR	XX	10-JUN-1998;	98US-0088734P
PR	XX	10-JUN-1998;	98US-0088740P
PR	XX	10-JUN-1998;	98US-0088811P
PR	XX	10-JUN-1998;	98US-0088824P
PR	XX	10-JUN-1998;	98US-0088825P
PR	XX	11-JUN-1998;	98US-0088863P
PR	XX	12-JUN-1998;	98US-0089105P
PR	XX	16-JUN-1998;	98US-0089514P
PR	XX	17-JUN-1998;	98US-0089653P
PR	XX	19-JUN-1998;	98US-0089952P
PR	XX	22-JUN-1998;	98US-0090246P
PR	XX	24-JUN-1998;	98US-0090444P
PR	XX	25-JUN-1998;	98US-0090668P
PR	XX	25-JUN-1998;	98US-0090686P
PR	XX	26-JUN-1998;	98US-0090862P
PR	XX	02-JUL-1998;	98US-0091628P
PR	XX	10-AUG-1998;	98US-0096012P
PR	XX	17-AUG-1998;	98US-0096757P
PR	XX	18-AUG-1998;	98US-0096949P
PR	XX	18-AUG-1998;	98US-0096959P
PR	XX	26-AUG-1998;	98US-0097954P
PR	XX	26-AUG-1998;	98US-0097971P
PR	XX	26-AUG-1998;	98US-0097979P
PR	XX	01-SEP-1998;	98US-0098749P
PR	XX	10-SEP-1998;	98US-0099741P
PR	XX	10-SEP-1998;	98US-0099763P
PR	XX	10-SEP-1998;	98US-0099792P
PR	XX	10-SEP-1998;	98US-0099812P
PR	XX	10-SEP-1998;	98US-0099815P
PR	XX	16-SEP-1998;	98US-0100627P
PR	XX	16-SEP-1998;	98US-0100629P
PR	XX	16-SEP-1998;	98US-0100630P
PR	XX	16-SEP-1998;	98US-0100683P
PR	XX	17-SEP-1998;	98US-0100684P
PR	XX	17-SEP-1998;	98US-0101738P
PR	XX	24-SEP-1998;	98US-0101743P
PR	XX	24-SEP-1998;	98US-0101916P
PR	XX	30-SEP-1998;	98US-0102570P
PR	XX	06-OCT-1998;	98US-0103449P
PR	XX	08-MAR-1999;	99US-00505028
PR	XX	14-MAY-1999;	99US-005010733
PR	XX	02-JUN-1999;	99US-005012252
PR	XX	01-SEP-1999;	99US-005020111
PR	XX	15-SEP-1999;	99US-005021194
PR	XX	15-SEP-1999;	99US-005021194
PR	XX	22-DEC-1999;	99US-005030720
PR	XX	18-FEB-2000;	2000US-00504341
PR	XX	18-FEB-2000;	2000US-00504342
PR	XX	22-FEB-2000;	2000US-005043414
PR	XX	01-MAR-2000;	2000US-00505601
PR	XX	30-MAR-2000;	2000US-00509439
PR	XX	22-MAY-2000;	2000US-005014022
PR	XX	02-JUN-2000;	2000US-005015264

23-AUG-2000; 2000MO-US023522.
PR 24-AUG-2000; 2000MO-US023328.
PR 10-NOV-2000; 2000MO-US030873.
PR 01-DEC-2000; 2000MO-US032378.
PR 20-DEC-2000; 2000MO-US034956.
PR 28-FEB-2001; 2001MO-US005520.
PR 01-MAR-2001; 2001MO-US006666.
PR 30-MAY-2001; 2001MO-US017443.
PR 01-JUN-2001; 2001MO-US017800.
PR 20-JUN-2001; 2001MO-US019692.
PR 29-JUN-2001; 2001MO-US021066.
PR 09-JUL-2001; 2001MO-US021735.
XX
XX
XX (GETH) GENENTECH INC.
PA
PI Saton DL, Filvarco E, Gerritsen ME, Goddard A, Godowski PJ;
PI Grimaldi JC, Gurney AL, Matanabe CK, Wood WI;
XX
XX WPI; 2002-731348/79.
DR N-PSDB; ABS74427.
XX
XX
XX New isolated secreted and transmembrane PRO polypeptide useful for
PT modulating biological activity of a cell, or for treating sports-related
PT joint problems, osteoarthritis or rheumatoid arthritis.
XX
XX
XX Claim 20; Fig 100; 399pp; English.
PS
XX
XX
XX The invention relates to an isolated secreted and transmembrane PRO
CC polypeptide having 80 % sequence identity to a sequence appearing as
CC ABG95851-ABG95934 or their associated signal peptide, or a sequence of an
CC extracellular domain of the proteins with their associated signal peptide
CC or lacking its associated signal peptide. Also included are the nucleic
CC acids encoding the proteins, vectors, host cells, fusion proteins and
CC antibodies which specifically bind to the proteins. The proteins are
CC useful for detecting a polypeptide designated as A, B, C or D in a sample
CC suspected of containing an A, B, C or D polypeptide, by contacting the
CC sample with a polypeptide designated as E, F, G, H or I (or vice versa)
CC and determining the formation of a A/E, B/F, B/G, C/H or D/I polypeptide
CC conjugate in the sample, where the formation of the conjugate is
CC indicative of the presence of an A, B, C or D polypeptide in the sample,
CC where A is a PRO10272 polypeptide, B is a PRO20110 polypeptide, C is a
CC PRO10096 polypeptide, D is a PRO19760 polypeptide, E is a PRO5801
CC polypeptide, F is a PRO1 polypeptide, G is a PRO20040 polypeptide, H is a
CC PRO20233 polypeptide and I is a PRO1890 polypeptide. The sample comprises
CC a cell suspected of expressing the A, B, C or D polypeptide. The E, F, G, H
CC or I polypeptide is labeled with a detectable label or is attached to a
CC solid support. The proteins are useful for linking a bioactive molecule
CC to a cell expressing a polypeptide designated as A, B, C or D or E, F, G,
CC H or I. The bioactive molecule is a toxin, a radiolabel or an antibody.
CC The bioactive molecule causes death of the cell. A, B, C, D, E, F, G, H
CC or I, or antibodies against them are useful for modulating a biological
CC activity of a cell expressing a polypeptide designated as A, B, C or D
CC E, F, G, H, or I. The cell is killed. The proteins are useful for
CC identifying agonists or antagonists, for the preparation of a medicament
CC useful in the treatment of a condition which is responsive to the
CC purposes, as molecular weight markers for protein electrophoresis
CC purposes, and as therapeutic agents for treating sports-related joint
CC problems, articular cartilage defects, osteoarthritis or rheumatoid
CC arthritis. Nucleic acids encoding the proteins are useful as
CC hybridization probes, in chromosome and gene mapping, in the generation
CC of anti-sense RNA and DNA, for the preparation of the proteins, to
CC generate transgenic or knockout animals which are useful in the
CC development and screening of therapeutic useful reagents, for chromosome
CC identification, and in gene therapy. The antibody is useful as a
CC therapeutic agent, in a diagnostic assay and for affinity purification of
CC the protein from recombinant cell culture natural sources. The present
CC sequence represents a novel secreted or transmembrane protein of the
CC invention
XX
XX
XX Sequence 596 AA;
SQ

6.3%; Score 180; DB 5; Length 596;
Best Local Similarity 19.2%; Pred NO. 0.00053;

```
Matches 109; Conservative 90; Mismatches 272; Indels 98; Gaps 19;
QY 4 GNLGNPNVNNISIPAPLPSCOTDAGAGCGQLINLTGPGSRALFTPVPRNSMADSGDNRA 63
Db 13 GILLHLERATNS-----NETSTSNANTGSSVLTSSG-----ASTATNSG 49
QY 64 SDVPLPVPNPRMLASEITLNDGFEVLHDHGLDITLNQIGSSV-FRVETQEDGKHIVAG 122
Db 50 SSTVSSGVSTATISGSSVTSN-GVSI-V-TNSEFFHTTSGISTAINSEFSTSSGSIATN 107
QY 123 QRMNVEYS--VVLSDQEVARLOSIDPEGKDFVTCGRGAGAHAVTAADITEARQRII 180
Db 108 SSSSTTSSGASTATNSE-----SSTPSGASVTYTNSSGSSVTSSGASTATNSSTVSSAS 163
QY 161 ELLEPKTGESKGAGBSKGVGELRNSNGAENTTQTQSTSTSSLRSDPKMLALGYAT 240
Db 164 TATNSESSTLSSGASTATN-SDSSTTSSGASTATNSSESSTSSG-----ASTAT 211
QY 241 GLIGLAANGIYQALALTPPEPDSPTTTDPDAASATETATRDQLTKFAQNPONOKVNI 300
Db 212 N-----SESTTSSRASATATNSSESTT--SSGASTATNSSEHTTNGAGTATNSSESTTS 264
QY 301 LG-----NAIPSGVLKDDVAVNIEEQAKAGEAEAKQQAENNAAQKKYDEQOAKR 351
Db 265 SGASTATNSDSSTVSSGA---STATNSSESTTSSGAST-----ATN 302
QY 352 QDELKVVSGAGYGL-SGALLCGGIGVAVTAAIHRKNDPEQOTTTTTTTTTSKRYEN 410
Db 303 SSSSTTSSGASTATNSDSSTTSSGAGTATNSSESTTSSGISTVYTNSESTPSSGANTATN 362
QY 411 KPANNTPAQGVNDPTGSEEDTMSRRSSMASTSTFPDSSIGTYONPVDVYTS----- 464
Db 363 SSSSTTSSGANTAT-NSSESTVSSGASTATNSSESTTSSGASTATNSSESTTSSGASTAT 421
QY 465 LHDQVPTNSNTSYQNMGNMTDSVYYSITIQHPRPDTTNDGARIAGNPAGIQSTYARIAL 524
Db 422 NSDSSTTSSSEASTATNSSESTTSSGISTVYTNSESTTSSGANTATNSGSSVTS----- 474
QY 525 SGGLRHDMGTLTGGSNSAVNTSNPPAPG 553
Db 475 AGSGTALTGHEHTTSSHA-STAVSEAKPG 502

RESULT 14
ABUS8554
ID ABUS8554 standard; protein; 596 AA.
XX
XX
AC ABUS8554;
XX
DT 15-APR-2003 (first entry)
XX
DE Human PRO polypeptide #155.
XX
XX Human; PRO: cytosolic; tumour; cancer; breast; lung; stomach; liver;
KW dog; cat; cow; horse; sheep; pig; goat; rabbit; ADERT;
KW antibody-dependent enzyme mediated prodng therapy.
XX
OS Homo sapiens.
XX
PN US2003027272-A1.
XX
PD 06-FEB-2003.
XX
PF 21-JUN-2002; 2002US-00176492.
XX
PR 18-SEP-1997; 97US-0059263P.
PR 18-SEP-1997; 97US-0059266P.
PR 17-OCT-1997; 97US-0062250P.
PR 21-OCT-1997; 97US-0063486P.
PR 24-OCT-1997; 97US-0063120P.
PR 24-OCT-1997; 97US-0063121P.
PR 28-OCT-1997; 97US-0063540P.
PR 28-OCT-1997; 97US-0063541P.
```

```
PR 28-OCT-1997; 97US-0063544P.
PR 28-OCT-1997; 97US-0063564P.
PR 29-OCT-1997; 97US-0063734P.
PR 31-OCT-1997; 97US-0063870P.
PR 31-OCT-1997; 97US-0064103P.
PR 13-NOV-1997; 97US-0065311P.
PR 21-NOV-1997; 97US-0066120P.
PR 24-NOV-1997; 97US-0066466P.
PR 24-NOV-1997; 97US-0066772P.
PR 11-DEC-1997; 97US-0069335P.
PR 12-DEC-1997; 97US-0069425P.
PR 17-DEC-1997; 97US-0069870P.
PR 18-DEC-1997; 97US-0068017P.
PR 10-MAR-1998; 98US-0077630P.
PR 11-MAR-1998; 98US-0077632P.
PR 11-MAR-1998; 98US-0077649P.
PR 20-MAR-1998; 98US-0078886P.
PR 20-MAR-1998; 98US-0078939P.
PR 27-MAR-1998; 98US-0079664P.
PR 27-MAR-1998; 98US-0079766P.
PR 31-MAR-1998; 98US-0080107P.
PR 31-MAR-1998; 98US-0080194P.
PR 01-APR-1998; 98US-0080327P.
PR 01-APR-1998; 98US-0080333P.
PR 08-APR-1998; 98US-0081049P.
PR 08-APR-1998; 98US-0081070P.
PR 09-APR-1998; 98US-0081195P.
PR 15-APR-1998; 98US-0081838P.
PR 21-APR-1998; 98US-0082568P.
PR 21-APR-1998; 98US-0082569P.
PR 22-APR-1998; 98US-0082704P.
PR 22-APR-1998; 98US-0082797P.
PR 28-APR-1998; 98US-0083322P.
PR 29-APR-1998; 98US-0083495P.
PR 29-APR-1998; 98US-0083496P.
PR 29-APR-1998; 98US-0083499P.
PR 29-APR-1998; 98US-0083559P.
PR 05-MAY-1998; 98US-0084366P.
PR 06-MAY-1998; 98US-0084414P.
PR 07-MAY-1998; 98US-0084639P.
PR 07-MAY-1998; 98US-0084640P.
PR 07-MAY-1998; 98US-0084643P.
PR 15-MAY-1998; 98US-0085579P.
PR 15-MAY-1998; 98US-0085580P.
PR 15-MAY-1998; 98US-0085582P.
PR 15-MAY-1998; 98US-0085700P.
PR 18-MAY-1998; 98US-0086023P.
PR 22-MAY-1998; 98US-0086392P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087208P.
PR 02-JUN-1998; 98US-0087609P.
PR 02-JUN-1998; 98US-0087759P.
PR 03-JUN-1998; 98US-0087827P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088032P.
PR 04-JUN-1998; 98US-0088033P.
PR 04-JUN-1998; 98US-0088326P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088213P.
PR 05-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088655P.
PR 10-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088811P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088825P.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088863P.
```

PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089598P.
PR 18-JUN-1998; 98US-0089653P.
PR 19-JUN-1998; 98US-0089908P.
PR 22-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090266P.
PR 22-JUN-1998; 98US-0090282P.
PR 24-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090680P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 26-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 01-JUL-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 02-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091632P.
PR 04-JUL-1998; 98US-0094006P.
PR 10-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095398P.
PR 17-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.
PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 26-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097852P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 01-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 02-SEP-1998; 98US-0098843P.
PR 09-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 15-SEP-1998; 98US-0099812P.
PR 16-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98US-0101751P.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.

PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 18-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102310P.
PR 29-SEP-1998; 98US-0102311P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.
PR 02-OCT-1998; 98US-0102687P.
PR 06-OCT-1998; 98US-0103258P.
PR 06-OCT-1998; 98US-0103449P.
PR 07-OCT-1998; 98US-00168978.

Query Match

6.3%; Score 180; DB 6; Length 596;

Best Local Similarity 19.2%; Pred. No. 0.00053;

Matches 109; Conservative 90; Mismatches 272; Indels 98; Gaps 19;

QY 4 GNLGNPNVNNISPPAPLPSCQTDGAGRGQLINSTGPLASRALFTPVNSMADSGDNRA 63
DB 13 GLHLHLEAATNS-----NETSTSNANTGSSVSSG-----ASTATNSG 49
QY 64 SDVPGLPVNPMRLAASEITLNDGFEVLHDGPIDLTNROIGSSV-PRVETQEDGKHIAVG 122
DB 50 SSVTSSGCVTATISGSSVTSN-GVSTV-TNSEPHITSSGISTATNSEFTASSGISITATN 107
QY 123 QNRGVETS--VVLSDQEVARLOSIDPEGKDFVTCGRGAGHAMVTVAADITEAROLL 180
DB 108 SSSSTTSSGASTATNSE-----STPSSGASTVTNSSGSSVTSAGASTATNSESSTVSRRS 163
QY 181 ELLEPKGTSGESKAGESKGVGLRESNCGAENTTETQTSSTSSLRSDPKLMLACTVAT 240
DB 164 TATNSESSTLSSGASTATN-SDSSTTSSGASTATNSESSTSSG-----ASTAT 211
QY 241 GLIGLAATGIVQALMTPEPDSPTTDPMAASATETATRDULTKEAPQNPQKNIDE 300
DB 212 N---SESSTVSSRASTATNSESSTT---SSGASTATNSESRTSNAGATATNSESSTTS 264
QY 301 LG-----NAIPSGVLKDDVVAIEEQAKAAGEAKQAQIENNAQAQKYDEQOAKR 351
DB 265 SGASTATNDSSTVSSGA---STAINSESSTTSSGAST-----ATN 302
QY 352 OEELKVSAGAYGL-SGALLILGGIGIVAVTAAALHRKNQPEQOTTTTTTTTTTARTVEN 410
DB 303 SSSSTTSSGASTATNDSSTTSSGAGTATNSESSTVSIGISTVTNSESSTPSSGANTATN 362
QY 411 KPAAINTPAGANDTPRSEPTMESRRSSMASTSTFTDTSIGTVQNPYADVITS----- 464
DB 363 SSSSTTSSGANTAT-NSSESSTVSSGASTATNSESSTTSSGASTATNSESSTTSSGASTAT 421
QY 465 LHDQVPTNSNTSVQNMGNMTDSVVYSTIQHPRDTTNDGALRLGNPSAGIOSTYARLAL 524
DB 422 NDSSTTSSSEASTATNSESSTVSIGISTVTNSESSTTSSGANTATNSSGSSVTS----- 474
QY 525 SCGLRHDMGCLTGGSNSAVNTSNPPAPG 553
DB 475 AGSGTAAITGMHTTSHA-STAVSEAKPG 502

RESULT 15

ABU88102
ID ABU88102 standard; protein; 596 AA.
XX
AC ABU88102;
XX
DT 07-JUL-2003 (first entry)
XX
DE Novel human secreted and transmembrane protein PRO1342.
XX
KW Human; secreted and transmembrane protein; PRO; gene therapy;
KW tumour necrosis factor-alpha release; TNF-alpha release;
KW chondrocyte proliferation; chondrocyte differentiation; tumour;
KW adrenal tumour; lung tumour; colon tumour; breast tumour;
KW prostate tumour; rectal tumour; cervical tumour; liver tumour.
XX
OS Homo sapiens.
XX
PN US2003032127-A1.
XX
PD 13-FEB-2003.
XX
PF 26-JUN-2002; 2002US-00183012.
XX
PR 18-SEP-1997; 97US-0059263P.
PR 18-SEP-1997; 97US-0059266P.
PR 17-OCT-1997; 97US-0062250P.
PR 21-OCT-1997; 97US-0063486P.
PR 24-OCT-1997; 97US-0063120P.
PR 24-OCT-1997; 97US-0063121P.
PR 28-OCT-1997; 97US-0063540P.
PR 28-OCT-1997; 97US-0063541P.
PR 28-OCT-1997; 97US-0063544P.
PR 28-OCT-1997; 97US-0063564P.
PR 29-OCT-1997; 97US-0063734P.
PR 31-OCT-1997; 97US-0063870P.
PR 31-OCT-1997; 97US-0064103P.
PR 13-NOV-1997; 97US-0065311P.
PR 21-NOV-1997; 97US-0066120P.
PR 24-NOV-1997; 97US-0066466P.
PR 24-NOV-1997; 97US-0066772P.
PR 11-DEC-1997; 97US-0069335P.
PR 12-DEC-1997; 97US-0069425P.
PR 17-DEC-1997; 97US-0069870P.
PR 18-DEC-1997; 97US-0068017P.
PR 10-MAR-1998; 98US-0077450P.
PR 11-MAR-1998; 98US-0077632P.
PR 11-MAR-1998; 98US-0077649P.
PR 20-MAR-1998; 98US-0078886P.
PR 20-MAR-1998; 98US-0078939P.
PR 27-MAR-1998; 98US-0079664P.
PR 27-MAR-1998; 98US-0079786P.
PR 31-MAR-1998; 98US-0080107P.
PR 31-MAR-1998; 98US-0080194P.
PR 01-APR-1998; 98US-0080327P.
PR 01-APR-1998; 98US-0080333P.
PR 08-APR-1998; 98US-0081049P.
PR 08-APR-1998; 98US-0081070P.
PR 09-APR-1998; 98US-0081195P.
PR 15-APR-1998; 98US-0081838P.
PR 21-APR-1998; 98US-0082568P.
PR 21-APR-1998; 98US-0082569P.
PR 22-APR-1998; 98US-0082704P.
PR 22-APR-1998; 98US-0083797P.
PR 28-APR-1998; 98US-0083322P.
PR 29-APR-1998; 98US-0083495P.
PR 29-APR-1998; 98US-0083496P.
PR 29-APR-1998; 98US-0083499P.
PR 29-APR-1998; 98US-0083559P.
PR 05-MAY-1998; 98US-0084366P.
PR 06-MAY-1998; 98US-0084414P.
PR 07-MAY-1998; 98US-0084639P.
PR 07-MAY-1998; 98US-0084640P.
PR 07-MAY-1998; 98US-0084643P.

PR 15-MAY-1998; 98US-0085579P.
PR 15-MAY-1998; 98US-0085580P.
PR 15-MAY-1998; 98US-0085582P.
PR 15-MAY-1998; 98US-0085700P.
PR 18-MAY-1998; 98US-0086023P.
PR 22-MAY-1998; 98US-0086392P.
PR 22-MAY-1998; 98US-0086486P.
PR 28-MAY-1998; 98US-0087098P.
PR 28-MAY-1998; 98US-0087208P.
PR 02-JUN-1998; 98US-0087609P.
PR 02-JUN-1998; 98US-0087759P.
PR 03-JUN-1998; 98US-0087827P.
PR 04-JUN-1998; 98US-0088025P.
PR 04-JUN-1998; 98US-0088028P.
PR 04-JUN-1998; 98US-0088029P.
PR 04-JUN-1998; 98US-0088033P.
PR 04-JUN-1998; 98US-0088326P.
PR 05-JUN-1998; 98US-0088167P.
PR 05-JUN-1998; 98US-0088202P.
PR 05-JUN-1998; 98US-0088212P.
PR 05-JUN-1998; 98US-0088217P.
PR 09-JUN-1998; 98US-0088655P.
PR 10-JUN-1998; 98US-0088722P.
PR 10-JUN-1998; 98US-0088738P.
PR 10-JUN-1998; 98US-0088740P.
PR 10-JUN-1998; 98US-0088811P.
PR 10-JUN-1998; 98US-0088824P.
PR 10-JUN-1998; 98US-0088825P.
PR 10-JUN-1998; 98US-0088826P.
PR 11-JUN-1998; 98US-0088861P.
PR 11-JUN-1998; 98US-0088863P.
PR 11-JUN-1998; 98US-0088876P.
PR 12-JUN-1998; 98US-0089090P.
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089312P.
PR 16-JUN-1998; 98US-0089314P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089599P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090252P.
PR 22-JUN-1998; 98US-0090254P.
PR 24-JUN-1998; 98US-0090433P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090676P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090689P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-00105413.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 26-JUN-1998; 98US-0091010P.
PR 01-JUL-1998; 98US-0091359P.
PR 01-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091632P.
PR 02-JUL-1998; 98US-0094006P.
PR 04-AUG-1998; 98US-0095282P.
PR 10-AUG-1998; 98US-0095988P.
PR 10-AUG-1998; 98US-0096012P.
PR 17-AUG-1998; 98US-0096757P.

```

PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0097022P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 01-SEP-1998; 98US-0098014P.
PR 01-SEP-1998; 98US-0098716P.
PR 02-SEP-1998; 98US-0098723P.
PR 02-SEP-1998; 98US-0098803P.
PR 02-SEP-1998; 98US-0098821P.
PR 09-SEP-1998; 98US-0098843P.
PR 10-SEP-1998; 98US-0099602P.
PR 10-SEP-1998; 98US-0099741P.
PR 10-SEP-1998; 98US-0099754P.
PR 10-SEP-1998; 98US-0099763P.
PR 10-SEP-1998; 98US-0099812P.
PR 15-SEP-1998; 98US-0100388P.
PR 16-SEP-1998; 98US-0100662P.
PR 16-SEP-1998; 98US-0100664P.
PR 16-SEP-1998; 98US-0101751P.
PR 16-SEP-1998; 98US-01019330.
PR 17-SEP-1998; 98US-0100683P.
PR 17-SEP-1998; 98US-0100684P.
PR 17-SEP-1998; 98US-0100919P.
PR 17-SEP-1998; 98US-0100930P.
PR 18-SEP-1998; 98US-0100849P.
PR 18-SEP-1998; 98US-0101014P.
PR 23-SEP-1998; 98US-0101068P.
PR 23-SEP-1998; 98US-0101471P.
PR 23-SEP-1998; 98US-0101472P.
PR 23-SEP-1998; 98US-0101475P.
PR 23-SEP-1998; 98US-0101477P.
PR 24-SEP-1998; 98US-0101738P.
PR 24-SEP-1998; 98US-0101739P.
PR 24-SEP-1998; 98US-0101743P.
PR 24-SEP-1998; 98US-0101922P.
PR 25-SEP-1998; 98US-0101786P.
PR 29-SEP-1998; 98US-0102207P.
PR 29-SEP-1998; 98US-0102240P.
PR 29-SEP-1998; 98US-0102330P.
PR 29-SEP-1998; 98US-0102331P.
PR 30-SEP-1998; 98US-0102487P.
PR 30-SEP-1998; 98US-0102570P.
PR 30-SEP-1998; 98US-0102571P.
PR 01-OCT-1998; 98US-0102684P.
PR 01-OCT-1998; 98US-0102687P.
PR 02-OCT-1998; 98US-0102965P.
PR 06-OCT-1998; 98US-0103258P.

```

Query Match 6.3%; Score 180; DB 6; Length 596;

Best Local Similarity 19.2%; Pred. No. 0.00053;

Matches 109; Conservative 90; Mismatches 272; Indels 98; Gaps 19;

```

QY 4 GNLGHNPNVNNNSPPAPPLPSOTDGCGRGQLNSTGPIGSRALFTPVNSMADSGDNRA 63
DB 13 GLLHLEAATNS-----NETSTISANTGSSVTSISG-----ASTATNSG 49
QY 64 SDVPGIPLVPMRLASETLTNDGFEVLADHGPLDLTNRQIGSSV-FRVETOEDGKHIAVG 122
DB 50 SSVTSGVGTATISGSSVTSN-GVSI-V-TNSEFHTTSSGISTATATNSEFTTASSGISIATN 107
QY 123 QNRNGVETS--VLSDOEFARLOSIDPEGKDFVFTGCRGAGHAMVTVASDITEARQRTL 180
DB 108 SSSSTTSSGASTATNSE---SSTPSSGASTVTNCGSVTSSGASTATNSESSVSSRAS 163
QY 181 ELLEPFGTGESKAGGSKGVGELRENSGAENKTETGTSTSTSLRSDPKMLALCTVAT 240

```

```

DB 164 TATNSESTSSGASTATN-SDSSTTSSGASTATNSESSTSSG-----ASTAT 211
QY 241 GLIGLATGIVQALATPPEDSPTTDPDAAASATETATDQUTKEAFONPDNQXNIDE 300
DB 212 N-----SESTTVSSRASTATNSESST---SGASTATNSESRTTSGAGATATNSESSTTS 264
QY 301 LG-----NAIPGVLKDVVNAIEOAKAAGEAKOALENNAAOKKYDEOQAKR 351
DB 265 SGASTATNDSSTVSSGA---STATNSESSTTSSGAST-----ATN 302
QY 352 QEELKTVSSGAGVGL-SCALILGGIGVAVTAALHRNQPVQETTTTTTTTTTARTVEN 410
DB 303 SESSSTTSSGASTATNDSSTTSSGAGATATNSESSTVSGISTVTNSESSTPSSGANTATN 362
QY 411 KPAANTPAQGNVUTPGESEDTMESRRSMASTSTFEDTSIGTVQNPYADVKTSS----- 464
DB 363 SESSSTTSSGANTPAT-NSESTTVSSGASTATNSESSTTSSGVSATNSESSTTSSGASTAT 421
QY 465 LHDQVPTNSNTSVQNMGTDSVYVSTIQHPRDRTDNGARLLGNPSAGIOSTYARLAL 524
DB 422 NSDSSTTSSSEASTATNSESSTVSSGISTVTNSESSTTSSGANTATNSGSSVTS----- 474
QY 525 SGGLRHDMGALTCGSNSAVNTSNPPAPG 553
DB 475 AGSGTALTGMTTSHSA-STAVSEAKPG 502

```

Search completed: August 1, 2006, 22:08:07

Job time : 203 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: August 1, 2006, 22:13:30 ; Search time 50 Seconds
(without alignments)
976.842 Million cell updates/sec

Title: US-09-189-415D-11

Perfect score: 2840

Sequence: 1 MFIGNLGHPNNVNSIPAP.....SNSAVNTSNNPAPSGHRRV 558

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 650591 seqs, 87530628 residues

Total number of hits satisfying chosen parameters: 650591

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Issued Patents AA:*
1: /EMC Celerra_SIDS3/ptodata/2/1aa/5_COMB.pep:*
2: /EMC Celerra_SIDS3/ptodata/2/1aa/6_COMB.pep:*
3: /EMC Celerra_SIDS3/ptodata/2/1aa/7_COMB.pep:*
4: /EMC Celerra_SIDS3/ptodata/2/1aa/H_COMB.pep:*
5: /EMC Celerra_SIDS3/ptodata/2/1aa/PCTUS_COMB.pep:*
6: /EMC Celerra_SIDS3/ptodata/2/1aa/RE_COMB.pep:*
7: /EMC Celerra_SIDS3/ptodata/2/1aa/backfilltest.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	180	6.3	596	2	US-10-012-231A-243 Sequence 243, App
2	180	6.3	596	2	US-10-015-389A-243 Sequence 243, App
3	180	6.3	596	2	US-10-006-768A-243 Sequence 243, App
4	180	6.3	596	2	US-10-015-671A-243 Sequence 243, App
5	180	6.3	596	2	US-10-015-393A-243 Sequence 243, App
6	180	6.3	596	2	US-10-011-833A-243 Sequence 243, App
7	180	6.3	596	2	US-10-006-041A-243 Sequence 243, App
8	180	6.3	596	2	US-10-012-064A-243 Sequence 243, App
9	180	6.3	596	2	US-10-015-392A-243 Sequence 243, App
10	180	6.3	596	3	US-10-011-795B-243 Sequence 243, App
11	180	6.3	596	3	US-10-015-386A-243 Sequence 243, App
12	180	6.3	596	3	US-10-012-121A-243 Sequence 243, App
13	180	6.3	596	3	US-10-006-485A-243 Sequence 243, App
14	180	6.3	596	3	US-10-006-746A-243 Sequence 243, App
15	180	6.3	596	3	US-10-012-752A-243 Sequence 243, App
16	180	6.3	596	3	US-10-017-253A-243 Sequence 243, App
17	180	6.3	596	3	US-10-017-519A-243 Sequence 243, App
18	180	6.3	596	3	US-10-015-715A-243 Sequence 243, App
19	180	6.3	596	3	US-10-007-236A-243 Sequence 243, App
20	172	6.1	2137	2	US-09-134-001C-4463 Sequence 4463, App
21	172	6.1	2137	2	US-09-252-991A-30227 Sequence 30227, A
22	170	6.0	529	2	US-09-248-796A-16703 Sequence 16703, A
23	169	6.0	2283	2	US-10-172-502-4 Sequence 4, Appl
24	166	5.8	2870	2	US-09-479-467A-15 Sequence 15, Appl
25	166	5.8	2870	2	US-09-655-160-15 Sequence 15, Appl
26	166	5.8	3178	2	US-09-479-467A-4 Sequence 4, Appl

27	166	5.8	3178	2	US-09-655-160-4	Sequence 4, Appl
28	164	5.8	1419	2	US-10-197-220-173	Sequence 173, App
29	163.5	5.8	821	2	US-09-556-877-195	Sequence 195, App
30	163.5	5.8	821	2	US-09-620-412C-195	Sequence 195, App
31	163.5	5.8	821	2	US-09-598-419-195	Sequence 195, App
32	163.5	5.8	1776	2	US-09-556-877-179	Sequence 179, App
33	163.5	5.8	1776	2	US-09-620-412C-179	Sequence 179, App
34	163.5	5.8	1776	2	US-09-598-419-179	Sequence 179, App
35	159.5	5.6	441	2	US-09-248-796A-15936	Sequence 15936, A
36	158	5.6	556	2	US-09-248-796A-17307	Sequence 17307, A
37	157.5	5.5	827	2	US-09-248-796A-17307	Sequence 17307, A
38	152.5	5.4	2045	2	US-09-268-347-49	Sequence 49, Appl
39	152	5.4	2314	2	US-10-369-983-16	Sequence 16, Appl
40	151.5	5.3	1154	3	US-09-538-092-330	Sequence 330, App
41	151	5.3	1306	2	US-09-316-048B-9	Sequence 9, Appl
42	150	5.3	1026	2	US-09-538-092-872	Sequence 872, App
43	149.5	5.3	785	2	US-08-046-585-5	Sequence 5, Appl
44	148	5.2	2035	1	US-08-393-703-5	Sequence 5, Appl
45	148	5.2	2035	1	US-08-393-703-5	Sequence 5, Appl

ALIGNMENTS

RESULT 1
US-10-012-231A-243
Sequence 243, Application US/10012231A
Patent No. 6924355
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan I.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE OF INVENTION: Acids Encoding the Same
FILE REFERENCE: P2830P1C23
CURRENT APPLICATION NUMBER: US/10/012,231A
CURRENT FILING DATE: 2002-06-10
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 477
SEQ ID NO 243
LENGTH: 596
TYPE: PRP
ORGANISM: Homo sapiens
US-10-012-231A-243
Query Match 6.3%; Score 180; DB 2; Length 596;
Best Local Similarity 19.2%; Pred. No. 1.2e-05;
Matches 109; Conservative 90; Mismatches 272; Indels 98; Gaps 19;
QY 4 GNLGNPNVNSIPAPLPISQTDGAGRGQLINSTPLGSRALFTPVNSMADSGDNA 63
DB 13 GLLHLEAATNS-----NETSANTGSSVISG-----ASTATNSG 49
QY 64 SVPGLPNPRLAASEITLNDGFEVLHDHGLDILNQIGSSV-FRVETQEDGHIWVG 122
DB 50 SSVTSVGSTATISSVTSN-GVSV-TNSEFTTSGISTATNSSEPTSSSGSISTATN 107
QY 123 QNRGVETS-VVLSDQEVARLQSIDPECKDKFVFTGRCGAGHAMVTVASDITTEARORIL 180
DB 108 SSSSTTSGASATATNSE-----SSTPSGASATVTVNSGSSVTSNGASATATNSSESTVSRA 163
QY 181 ELLEPKGTGSGKAGSGKGVGELRBSNGAENTTETOTSTSTSSLRDPKMLALGTAVAT 240

```
Db      164 TATNSESSTLSSGASTATN-SDSSTTSSGASTATNSESSTTSSG-----ASTAT 211
Qy      241 GLIGLAAGIYOALALTEPPDSPPTTDPDAASATETATROLTEAFONPNQVNI 300
Db      212 N-----SESTTVSSRASTATNSESSTT---SSGASTATNSESSTRTNGAGTATNSESSTTS 264
Qy      301 LG-----NAIPSGVLKDDVANIIEQAKAAGEAKQQAIEENNAQKXYDQAKR 351
Db      265 SGASTATNDSSTVSSGA---STATNSESSTTSSGAST-----ATN 302
Qy      352 QEELKVSSGAGYGL-SGALLIGGIGVAVTAALHRKNQPVQOTTTTTTTTTTSARVEN 410
Db      303 SESSTTSSGASTATNDSSTTSSGAGTATNSESSTVSGISTVTNSESSTSPSGANTATN 362
Qy      411 KPANTPAQGNVDTGSGEDTMRSSMASTSTFFDTSIGTVONPYADVKT----- 464
Db      363 SESSTTSSGANTAT-NSBSTTVSSGASTATNSESSTTSSGASTATNSESSTTSSGASTAT 421
Qy      465 LHDQVPTSNSTSVONKGNNTDSVYVSTIOHPPDITDNGARLLGNPSAGIOSTYARLAL 524
Db      422 NSDSTTSSGASTATNSESSTVSSGISTVTNSESSTTSSGANTATNSSSVTS----- 474
Qy      525 SGGLRHDMGGLTGGSNSAVNTSNNPPAPG 553
Db      475 AGSGTAAITGMHTTSHSA-STAVSEAKPG 502
```

RESULT 2
US-10-015-389A-243
Sequence 243, Application US/10015389A

Patent No. 6936436
GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Baton, Dan I.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
TITLE OF INVENTION: Acids Encoding the Same
FILE REFERENCE: P2830P1C48
CURRENT APPLICATION NUMBER: US/10/015,389A
CURRENT FILING DATE: 2002-06-25
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 477
SEQ ID NO 243
LENGTH: 596
TYPE: PRT
ORGANISM: Homo sapiens
US-10-015-389A-243

Query Match 6.3%; Score 180; DB 2; Length 596;
Best Local Similarity 19.2%; Pred. No. 1.2e-05;

Matches 109; Conservative 90; Mismatches 272; Indels 98; Gaps 19;

```
Qy      4 GNLGNPNVNNISIPAPLPISQTDGAGRGQLINSTPGISRALFTPVNSMADSGDNR 63
Db      13 GILLHLERATNS-----NETSTSNANTGSSVSSG-----ASTATNSG 49
Qy      64 SDVPGLPVNPMRLAASEITLNDGEVLHDHGPLDTLNRQIGSSV-FRVETOEDGKHIAVG 122
Db      50 SSVTSAGVSTATISGSSVTSN-GVSTV-TNSEPHTTSSGISTATNSEFSTASSGISTATN 107
Qy      123 QNNGVETS--VVLSDQEVARLQSIDPECKDKFVFTGGRGAGHAMVTASDITEARQRL 180
```

```
Db      108 SESSTTSSGASTATNSE-----STTPSSGASTVTVNSGSSVTSAGASTATNSESSTVSSRAS 163
Qy      181 ELLEPKTGKESKAGESKGVLEESNGAGANTETETOTISTSSLRSPKXLMALGVAT 240
Db      164 TATNSESSTLSSGASTATN-SDSSTTSSGASTATNSESSTTSSG-----ASTAT 211
Qy      241 GLIGLAAGIYOALALTEPPDSPPTTDPDAASATETATROLTEAFONPNQVNI 300
Db      212 N-----SESTTVSSRASTATNSESSTT---SSGASTATNSESSTRTNGAGTATNSESSTTS 264
Qy      301 LG-----NAIPSGVLKDDVANIIEQAKAAGEAKQQAIEENNAQKXYDQAKR 351
Db      265 SGASTATNDSSTVSSGA---STATNSESSTTSSGAST-----ATN 302
Qy      352 QEELKVSSGAGYGL-SGALLIGGIGVAVTAALHRKNQPVQOTTTTTTTTTTSARVEN 410
Db      303 SESSTTSSGASTATNDSSTTSSGAGTATNSESSTVSSGISTVTNSESSTSPSGANTATN 362
Qy      411 KPANTPAQGNVDTGSGEDTMRSSMASTSTFFDTSIGTVONPYADVKT----- 464
Db      363 SESSTTSSGANTAT-NSBSTTVSSGASTATNSESSTTSSGASTATNSESSTTSSGASTAT 421
Qy      465 LHDQVPTSNSTSVONKGNNTDSVYVSTIOHPPDITDNGARLLGNPSAGIOSTYARLAL 524
Db      422 NSDSTTSSGASTATNSESSTVSSGISTVTNSESSTTSSGANTATNSSSVTS----- 474
Qy      525 SGGLRHDMGGLTGGSNSAVNTSNNPPAPG 553
Db      475 AGSGTAAITGMHTTSHSA-STAVSEAKPG 502
```

RESULT 3
US-10-006-768A-243
Sequence 243, Application US/10006768A

Patent No. 6936697
GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Baton, Dan I.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
TITLE OF INVENTION: Acids Encoding the Same
FILE REFERENCE: P2830P1C10
CURRENT APPLICATION NUMBER: US/10/006,768A
CURRENT FILING DATE: 2002-03-05
NUMBER OF SEQ ID NOS: 477
Prior Application removed - See File Wrapper or Palm
SEQ ID NO 243
LENGTH: 596
TYPE: PRT
ORGANISM: Homo sapiens
US-10-006-768A-243

Query Match 6.3%; Score 180; DB 2; Length 596;
Best Local Similarity 19.2%; Pred. No. 1.2e-05;

Matches 109; Conservative 90; Mismatches 272; Indels 98; Gaps 19;

```
Qy      4 GNLGNPNVNNISIPAPLPISQTDGAGRGQLINSTPGISRALFTPVNSMADSGDNR 63
Db      13 GILLHLERATNS-----NETSTSNANTGSSVSSG-----ASTATNSG 49
Qy      64 SDVPGLPVNPMRLAASEITLNDGEVLHDHGPLDTLNRQIGSSV-FRVETOEDGKHIAVG 122
```

Db 50 SSVTSSGVSATISGSSVTSN-GVSIY-TNSEPHITSSGISTATNSEFTASSGISIATN 107
QY 123 QRMGVETS--VLSDDEYALQSIDPEGKDKFVFTGRGAGHAMTVASDITEARQIL 180
Db 108 SSSSTSSGASTATNSE---SSTPSSGASTVNTSSSVTS SGASTATNSESSVSSRAS 163
QY 181 ELLEPKGTGSKAGGSKGVGELRESNCGAENTTETOTSTSSLRSDPRLMALGTVA 240
Db 164 TATNSESSSTSSGASTATN-SDSSTSSGASTATNSESSSTSSG-----ASTAT 211
QY 241 GLIGLAATGIVQALALTPEDSPFTTDPDAASATETATRDOLTKAFOQNDQKVNIDE 300
Db 212 N-----SSSTVSSRASATATNSESTT---SSGASTATNSESRRTSSNAGATATNSESTTS 264
QY 301 LG-----NAIPSGVLKDDVVAIEQAKAAGEAKQAIEENNAQAKKYDEQOAKR 351
Db 265 SGASTATNSDSSVSSGA---STATNSESTSSGAST-----ATN 302
QY 352 QEELKTVSSGAGYGL-SCALILGGIGVAVTAALHRKNQVPEQTTTTTTTTTTSARTVEN 410
Db 303 SSSSTSSGASTATNSDSSSTSSGAGTATNSESTVSSGISTVNTNSESTSSGANTATN 362
QY 411 KPANNTPAQGNVDTPPSEDTMESRRSSMASTSTFPDTSSIGTVQNPYADVKT-----464
Db 363 SSSSTSSGANTAT-NSESTVSSGASTATNSESTSSGISTVNTNSESTSSGANTATN 421
QY 465 LHDQVPTNSNTSVNMGMNTDSVYSTIOHPRDITDNGARLLGPSAGIOSTVARLAL 524
Db 422 NSDSSTSSSEASTATNSESTVSSGISTVNTNSESTSSGANTATNSGSVTS-----474
QY 525 SGGLRHDGGLTGGSNSAVTNSNPPAPG 553
Db 475 AGSGTALTGMHTTSSHA-STAVSEAKPG 502

RESULT 4
US-10-015-671A-243
; Sequence 243, Application US/10015671A
; Patent No. 6946263
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnovers, Luc
; APPLICANT: Eaton, Dan I.
; APPLICANT: Feirara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2830P1C47
; CURRENT APPLICATION NUMBER: US/10/015,671A
; CURRENT FILING DATE: 2001-12-11
; Prior application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 477
; SEQ ID NO 243
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-015-671A-243

Query Match 6.3%; Score 180; DB 2; Length 596;
Best Local Similarity 19.2%; Pred. No. 1.2e-05;
Matches 109; Conservative 90; Mismatches 272; Indels 98; Gaps 19;
QY 4 GNLGHRPNVNNISPPAPLPISQITDAGGGRGLINSIGPLSGRALLFPVNSMADSGDNRA 63

Db 13 GULLHLEAATNS-----NETSTANTGSSVISSG-----ASTATNSG 49
QY 64 SDVPGI-PVNEPRLAASEITLNDGFEVLHDGPIDLTLNRQIGSSV-FRVEFOEDQKHIAVG 122
Db 50 SSVTSSGVSATISGSSVTSN-GVSIY-TNSEPHITSSGISTATNSEFTASSGISIATN 107
QY 123 QRMGVETS--VLSDDEYALQSIDPEGKDKFVFTGRGAGHAMTVASDITEARQIL 180
Db 108 SSSSTSSGASTATNSE---SSTPSSGASTVNTSSSVTS SGASTATNSESSVSSRAS 163
QY 181 ELLEPKGTGSKAGGSKGVGELRESNCGAENTTETOTSTSSLRSDPRLMALGTVA 240
Db 164 TATNSESSSTSSGASTATN-SDSSTSSGASTATNSESSSTSSG-----ASTAT 211
QY 241 GLIGLAATGIVQALALTPEDSPFTTDPDAASATETATRDOLTKAFOQNDQKVNIDE 300
Db 212 N-----SSSTVSSRASATATNSESTT---SSGASTATNSESRRTSSNAGATATNSESTTS 264
QY 301 LG-----NAIPSGVLKDDVVAIEQAKAAGEAKQAIEENNAQAKKYDEQOAKR 351
Db 265 SGASTATNSDSSVSSGA---STATNSESTSSGAST-----ATN 302
QY 352 QEELKTVSSGAGYGL-SCALILGGIGVAVTAALHRKNQVPEQTTTTTTTTTTSARTVEN 410
Db 303 SSSSTSSGASTATNSDSSSTSSGAGTATNSESTVSSGISTVNTNSESTSSGANTATN 362
QY 411 KPANNTPAQGNVDTPPSEDTMESRRSSMASTSTFPDTSSIGTVQNPYADVKT-----464
Db 363 SSSSTSSGANTAT-NSESTVSSGASTATNSESTSSGISTVNTNSESTSSGANTATN 421
QY 465 LHDQVPTNSNTSVNMGMNTDSVYSTIOHPRDITDNGARLLGPSAGIOSTVARLAL 524
Db 422 NSDSSTSSSEASTATNSESTVSSGISTVNTNSESTSSGANTATNSGSVTS-----474
QY 525 SGGLRHDGGLTGGSNSAVTNSNPPAPG 553
Db 475 AGSGTALTGMHTTSSHA-STAVSEAKPG 502

RESULT 5
US-10-015-393A-243
; Sequence 243, Application US/10015393A
; Patent No. 6951737
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnovers, Luc
; APPLICANT: Eaton, Dan I.
; APPLICANT: Feirara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2830P1C46
; CURRENT APPLICATION NUMBER: US/10/015,393A
; CURRENT FILING DATE: 2002-06-10
; Prior application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 477
; SEQ ID NO 243
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-015-393A-243

Query Match 6.3%; Score 180; DB 2; Length 596;

Best Local Similarity 19.2%; Pred. No. 1.2e-05;
Matches 109; Conservative 90; Mismatches 272; Indels 98; Gaps 19;

```
QY 4 GNLGNPNVNNISIPAPLPSPOTDAGRGQLINSTGLPSRALFTPVNRMADSGDNRA 63
Db 13 GLLHLLEAATNS-----NETSTANTGSSVSSG-----ASTATNSG 49
QY 64 SDVPLPVNPKRLAASEITLNDGFEVLHDHPLDPLTNQIGSSV-FRVETOEDGKHIAVG 122
Db 50 SVSTSSGVSTATISGSSVTSN-GVSIY-TNSEPHITSSGISTATNSEFSTASSGISIATN 107
QY 123 QNNGVETS--VVLSDQEFARLQSIDPEKDKFVFTGRCGAGHAMVYASDITEARQIL 180
Db 108 SSSSTSSGASTATNSE-----SSTPSSGASTVYTNSSGSSVTSNGASTATNSESTVSSRAS 163
QY 181 ELLEPKGTGSGAGESKGVGELRESNGAENTTETQSTSTSLRSDPKMLALGTAT 240
Db 164 TATNSESSTLSSGASTATN-SDSSTSSGASTATNSESTTSSG-----ASTAT 211
QY 241 GLIGLAATGIVQALALTPEDPSPTTDDPAAASATETATRDQLTKEAFQNPNOXNIDE 300
Db 212 N-----SESTVSSRASTATNSESSTT---SSGASTATNSESRTTNGAGTATNSESSTTS 264
QY 301 LG-----NALPSGVLKDDVANIEEQAKAAGEAKQQAIEENNAQAKKYDEQAKR 351
Db 265 SGASTATNDSSTVSSGA---STATNSESSTSSGAST-----ATN 302
QY 352 QEELVSSGAGYGL-SCALILGGIGVAVTALHKNQVPEQTITTTTTTTTSARVYEN 410
Db 303 SSSSTSSGASTATNDSSTSSGAGTATNSESSTVSSGISTVYTNSESSTSSGANTATN 362
QY 411 KPANTPAQGNVDTPGSEDTMESRRSSMASTSTFFDTSSIGTVONPYADVTS----- 464
Db 363 SSSSTSSGANTAT-NSBSTVSSGASTATNSESSTSSGCVSTATNSESSTSSGASTAT 421
QY 465 LHDQVPTNSNTSVQNMGNITDSVYVSTIQHPPRDTTNGARLLGNPSAGIOSTYARLAL 524
Db 422 NSDSSTTSEASTATNSESSTVSSGISTVYTNSESSTSSGANTATNSSGSVTS----- 474
QY 525 SGGLRHDMGGLTGGSNSAVNTSNNPPAPG 553
Db 475 AGSGTALTLGKMTTSHSA-STAVSEAKPG 502
```

RESULT 6

US-10-011-833A-243
Sequence 243, Application US/10011833A

Patent No. 6951920

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.

APPLICANT: Botstein, David

APPLICANT: Desnoyers, Luc

APPLICANT: Baton, Dan I.

APPLICANT: Ferrara, Napoleone

APPLICANT: Fong, Sherman

APPLICANT: Gao, Wei-Qiang

APPLICANT: Goddard, Audrey

APPLICANT: Godowski, Paul J.

APPLICANT: Grimaldi, Christopher J.

APPLICANT: Gurney, Austin L.

APPLICANT: Hillan, Kenneth J.

APPLICANT: Pan, James

APPLICANT: Paoni, Nicholas F.

TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

FILE REFERENCE: P2830PIC22

CURRENT APPLICATION NUMBER: US/10/011,833A

CURRENT FILING DATE: 2002-06-25

Prior Application removed - See File Wrapper or Palm

SEQU ID NO 243

LENGTH: 596

TYPE: PRT

ORGANISM: Homo sapiens
US-10-011-833A-243

Query Match 6.3%; Score 180; DB 2; Length 596;

Best Local Similarity 19.2%; Pred. No. 1.2e-05;

Matches 109; Conservative 90; Mismatches 272; Indels 98; Gaps 19;

```
QY 4 GNLGNPNVNNISIPAPLPSPOTDAGRGQLINSTGLPSRALFTPVNRMADSGDNRA 63
Db 13 GLLHLLEAATNS-----NETSTANTGSSVSSG-----ASTATNSG 49
QY 64 SDVPLPVNPKRLAASEITLNDGFEVLHDHPLDPLTNQIGSSV-FRVETOEDGKHIAVG 122
Db 50 SVSTSSGVSTATISGSSVTSN-GVSIY-TNSEPHITSSGISTATNSEFSTASSGISIATN 107
QY 123 QNNGVETS--VVLSDQEFARLQSIDPEKDKFVFTGRCGAGHAMVYASDITEARQIL 180
Db 108 SSSSTSSGASTATNSE-----SSTPSSGASTVYTNSSGSSVTSNGASTATNSESTVSSRAS 163
QY 181 ELLEPKGTGSGAGESKGVGELRESNGAENTTETQSTSTSLRSDPKMLALGTAT 240
Db 164 TATNSESSTLSSGASTATN-SDSSTSSGASTATNSESSTTSSG-----ASTAT 211
QY 241 GLIGLAATGIVQALALTPEDPSPTTDDPAAASATETATRDQLTKEAFQNPNOXNIDE 300
Db 212 N-----SESTVSSRASTATNSESSTT---SSGASTATNSESRTTNGAGTATNSESSTTS 264
QY 301 LG-----NALPSGVLKDDVANIEEQAKAAGEAKQQAIEENNAQAKKYDEQAKR 351
Db 265 SGASTATNDSSTVSSGA---STATNSESSTSSGAST-----ATN 302
QY 352 QEELVSSGAGYGL-SCALILGGIGVAVTALHKNQVPEQTITTTTTTTTSARVYEN 410
Db 303 SSSSTSSGASTATNDSSTSSGAGTATNSESSTVSSGISTVYTNSESSTSSGANTATN 362
QY 411 KPANTPAQGNVDTPGSEDTMESRRSSMASTSTFFDTSSIGTVONPYADVTS----- 464
Db 363 SSSSTSSGANTAT-NSBSTVSSGASTATNSESSTSSGCVSTATNSESSTSSGASTAT 421
QY 465 LHDQVPTNSNTSVQNMGNITDSVYVSTIQHPPRDTTNGARLLGNPSAGIOSTYARLAL 524
Db 422 NSDSSTTSEASTATNSESSTVSSGISTVYTNSESSTSSGANTATNSSGSVTS----- 474
QY 525 SGGLRHDMGGLTGGSNSAVNTSNNPPAPG 553
Db 475 AGSGTALTLGKMTTSHSA-STAVSEAKPG 502
```

RESULT 7

US-10-006-041A-243
Sequence 243, Application US/10006041A

Patent No. 6951921

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.

APPLICANT: Botstein, David

APPLICANT: Desnoyers, Luc

APPLICANT: Baton, Dan I.

APPLICANT: Ferrara, Napoleone

APPLICANT: Fong, Sherman

APPLICANT: Gao, Wei-Qiang

APPLICANT: Goddard, Audrey

APPLICANT: Godowski, Paul J.

APPLICANT: Grimaldi, Christopher J.

APPLICANT: Gurney, Austin L.

APPLICANT: Hillan, Kenneth J.

APPLICANT: Pan, James

APPLICANT: Paoni, Nicholas F.

TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

FILE REFERENCE: P2830PIC8

CURRENT APPLICATION NUMBER: US/10/006,041A

CURRENT FILING DATE: 2001-12-06

Prior Application removed - See File Wrapper or Palm

SEQU ID NO 243

LENGTH: 596

TYPE: PRT

NUMBER OF SEQ ID NOS: 477
SEQ ID NO 243
LENGTH: 596
TYPE: PRT
ORGANISM: Homo sapiens
US-10-006-041A-243

Query Match 6.3%; Score 180; DB 2; Length 596;
Best Local Similarity 19.2%; Pred. No. 1.2e-05;
Matches 109; Conservative 90; Mismatches 272; Indels 98; Gaps 19;

QY 4 GNLGHPNPNNSIPAPPLPSQTDGAGRGQLINSTGPLSRALFPVRNSMADSGDNRA 63
DB 13 GLLLHLEAATNS-----NETSTANTGSSVSSG-----ASTATNG 49
QY 64 SDVPLGPNPMLAASEITLNDGFEVLHDHGPLDLNROISSV-FRVETQEDGKHIAVG 122
DB 50 SSVTSSGVSSTATISGSSVTSN-GVSIY-TNSEPHHTSSGISTATNSEFTASSGISTATN 107
QY 123 QRGVETS--VILSDQEVARLQSIDPEGKDKFVTGGRGAGAMTVASDITEARQRL 180
DB 108 SESSSTSSGASTATNSE---SSTPSSGASTVTNNGSSVTSNGASTATNSESSVSSRAS 163
QY 181 ELLEPKGTGSKAGSKGVGELRESNGAENTTETQSTSTSLRSDPKMLALGTVA 240
DB 164 TATNSESSTLSSGASTATN-SDSSTSSGASTATNSESSSTSSG-----ASTAT 211
QY 241 GLIGLAATGIVQALALTPEDSPFTTDPDAAASATETATRDQTKFAFQNPQKNVIDE 300
DB 212 N-----SESSVSSRASATATNSESTT---SSGASTATNSESRITTSNGAGTATNSESSSTS 264
QY 301 LG-----NAIPSGVLKDVVANIEBOAKAAGEAKQOALENNAOAKKYDEQAKR 351
DB 265 SGASTATNDSSTVSSGA---STATNSESTSSGAST-----ATN 302
QY 352 QEELKXSSGAGYGL-SCALLIGGIGVAVTAALHRKQVPEQTTTTTTTTTTSARTVEN 410
DB 303 SESSSTSSGASTATNDSSTSSGAGTATNSESSVSSGISTYTNSSSTPSSGANTATN 362
QY 411 KPANNTPAQGNVDTPESEDPTMESRRSSMASTSTFPDTSIGTVQNPYADVKT 464
DB 363 SESSSTSSGANTAT-NSSESTVSSGASTATNSESSSTSSGVTATNSESSSTSSGASTAT 421
QY 465 LHDQVPTNSNTSVQNMGTDSVYSTIQHPROTDTNARLLGNPSAGIOGTYARLAL 524
DB 422 NSDSSTSSSEASATATNSESSVSSGISTVYNSESSSTSSGANTATNNGSSVTS----- 474
QY 525 SCGLRHDGCLTGSNSAVTNSNPPAPG 553
DB 475 AGSGTALTGMTTSHSA-STAVSEAKPG 502

RESULT 8
US-10-012-064A-243
Sequence 243, Application US/10012064A
Patent No. 6953841
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Baton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
TITLE OF INVENTION: Acids Encoding the Same

FILE REFERENCE: P2830P1C19
CURRENT APPLICATION NUMBER: US/10/012,064A
CURRENT FILING DATE: 2002-07-15
PRIOR APPLICATION NUMBER: 60/098716
PRIOR FILING DATE: 1998-09-01
PRIOR APPLICATION NUMBER: 60/098723
PRIOR FILING DATE: 1998-09-01
PRIOR APPLICATION NUMBER: 60/098749
PRIOR FILING DATE: 1998-09-01
PRIOR APPLICATION NUMBER: 60/098750
PRIOR FILING DATE: 1998-09-01
PRIOR APPLICATION NUMBER: 60/098803
PRIOR FILING DATE: 1998-09-02
PRIOR APPLICATION NUMBER: 60/098821
PRIOR FILING DATE: 1998-09-02
PRIOR APPLICATION NUMBER: 60/098843
PRIOR FILING DATE: 1998-09-02
PRIOR APPLICATION NUMBER: 60/099536
PRIOR FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: 60/099596
PRIOR FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: 60/099598
PRIOR FILING DATE: 1998-09-09
Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 477
SEQ ID NO 243
LENGTH: 596
TYPE: PRT
ORGANISM: Homo sapiens
US-10-012-064A-243

Query Match 6.3%; Score 180; DB 2; Length 596;
Best Local Similarity 19.2%; Pred. No. 1.2e-05;
Matches 109; Conservative 90; Mismatches 272; Indels 98; Gaps 19;

QY 4 GNLGHPNPNNSIPAPPLPSQTDGAGRGQLINSTGPLSRALFPVRNSMADSGDNRA 63
DB 13 GLLLHLEAATNS-----NETSTANTGSSVSSG-----ASTATNG 49
QY 64 SDVPLGPNPMLAASEITLNDGFEVLHDHGPLDLNROISSV-FRVETQEDGKHIAVG 122
DB 50 SSVTSSGVSSTATISGSSVTSN-GVSIY-TNSEPHHTSSGISTATNSEFTASSGISTATN 107
QY 123 QRGVETS--VILSDQEVARLQSIDPEGKDKFVTGGRGAGAMTVASDITEARQRL 180
DB 108 SESSSTSSGASTATNSE---SSTPSSGASTVTNNGSSVTSNGASTATNSESSVSSRAS 163
QY 181 ELLEPKGTGSKAGSKGVGELRESNGAENTTETQSTSTSLRSDPKMLALGTVA 240
DB 164 TATNSESSTLSSGASTATN-SDSSTSSGASTATNSESSSTSSG-----ASTAT 211
QY 241 GLIGLAATGIVQALALTPEDSPFTTDPDAAASATETATRDQTKFAFQNPQKNVIDE 300
DB 212 N-----SESSVSSRASATATNSESTT---SSGASTATNSESRITTSNGAGTATNSESSSTS 264
QY 301 LG-----NAIPSGVLKDVVANIEBOAKAAGEAKQOALENNAOAKKYDEQAKR 351
DB 265 SGASTATNDSSTVSSGA---STATNSESTSSGAST-----ATN 302
QY 352 QEELKXSSGAGYGL-SCALLIGGIGVAVTAALHRKQVPEQTTTTTTTTTTSARTVEN 410
DB 303 SESSSTSSGASTATNDSSTSSGAGTATNSESSVSSGISTYTNSSSTPSSGANTATN 362
QY 411 KPANNTPAQGNVDTPESEDPTMESRRSSMASTSTFPDTSIGTVQNPYADVKT 464
DB 363 SESSSTSSGANTAT-NSSESTVSSGASTATNSESSSTSSGVTATNSESSSTSSGASTAT 421
QY 465 LHDQVPTNSNTSVQNMGTDSVYSTIQHPROTDTNARLLGNPSAGIOGTYARLAL 524
DB 422 NSDSSTSSSEASATATNSESSVSSGISTVYNSESSSTSSGANTATNNGSSVTS----- 474
QY 525 SCGLRHDGCLTGSNSAVTNSNPPAPG 553

Db 475 AGSGTAALTGMHTTSSHA-STAVSEAKPG 502

RESULT 9

US-10-015-392A-243

/ Sequence 243, Application US/10015392A

/ Patent No. 6972186

/ GENERAL INFORMATION:

/ APPLICANT: Baker, Kevin P.

/ APPLICANT: Botstein, David

/ APPLICANT: Desmoyers, Luc

/ APPLICANT: Baton, Dan I.

/ APPLICANT: Ferrara, Napoleone

/ APPLICANT: Fong, Sherman

/ APPLICANT: Gao, Wei-Qiang

/ APPLICANT: Goddard, Audrey

/ APPLICANT: Godowski, Paul J.

/ APPLICANT: Grimaldi, Christopher J.

/ APPLICANT: Gurney, Austin L.

/ APPLICANT: Hillan, Kenneth J.

/ APPLICANT: Pan, James

/ APPLICANT: Paoni, Nicholas F.

/ TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

/ FILE REFERENCE: P2830P1C58

/ CURRENT FILING DATE: 2001-12-12

/ PRIOR FILING DATE: 1998-09-01

/ PRIOR APPLICATION NUMBER: 60/098716

/ PRIOR FILING DATE: 1998-09-01

/ PRIOR APPLICATION NUMBER: 60/098723

/ PRIOR FILING DATE: 1998-09-01

/ PRIOR APPLICATION NUMBER: 60/098749

/ PRIOR FILING DATE: 1998-09-01

/ PRIOR APPLICATION NUMBER: 60/098750

/ PRIOR FILING DATE: 1998-09-01

/ PRIOR APPLICATION NUMBER: 60/098803

/ PRIOR FILING DATE: 1998-09-02

/ PRIOR APPLICATION NUMBER: 60/098821

/ PRIOR FILING DATE: 1998-09-02

/ PRIOR APPLICATION NUMBER: 60/098843

/ PRIOR FILING DATE: 1998-09-02

/ PRIOR APPLICATION NUMBER: 60/099536

/ PRIOR FILING DATE: 1998-09-09

/ PRIOR APPLICATION NUMBER: 60/099596

/ PRIOR FILING DATE: 1998-09-09

/ PRIOR APPLICATION NUMBER: 60/099598

/ PRIOR FILING DATE: 1998-09-09

/ Remaining Prior Application data removed - See file wrapper or PALM.

/ NUMBER OF SEQ ID NOS: 477

/ SEQ ID NO 243

/ LENGTH: 596

/ TYPE: PRT

/ ORGANISM: Homo sapiens

US-10-015-392A-243

Query Match 6.3%; Score 180; DB 2; Length 596;

Best Local Similarity 19.2%; Pred. No. 1.2e-05;

Matches 109; Conservative 90; Mismatches 272; Indels 98; Gaps 19;

Qy 4 GNLGNPNVNSIPAPPLPSQTDGAGRGOLINSTGPLGRALFTPVNSNADSGDNRA 63

Db 13 GLLHLLEAATNS-----NETSNTANGSSVISG-----ASTATNG 49

Qy 64 SVPLGLPVNPMPLAASEITLNDGFEVLHDHGPLDTLNRIGSSV-PRVETQEDKHIANG 122

Db 50 SSVTSSGSTATISGSSVSN-GVSTV-TNSEFHTTSGISTATNSEFSTASSGSIATN 107

Qy 123 QHNGVETS--VVISDOEVARLQSIDPEGKDFVTGKGGAGHAMVTASDITTEARQRL 180

Db 108 SSSSTTSSGASTATNSE---STPSSGASTYTNSSSVTSSGASTATNSESSYSSRAS 163

Qy 181 ELLEPKGTGESHKAGSGKGVGELRESNSGAENTTETOTJSTSSLSRDPKMLALGTVA 240

Db 164 TATNSESSTLSSGASTATN-SDSSTTSSGASTATNSESSTTSSG-----ASTAT 211

Qy 241 GLIGLAATGIYQALALTEPPDSPTTDPDAAASATEPTARQQLTEAFQPNQNKVNIDE 300

Db 212 N-----SESTVSSSRSTATNSESSTT---SSGASTATNSERTTNSGATATNSESSTTS 264

Qy 301 LG-----NALPFGVLKDDVVANIEQAKAAGEAKQQA1ENNAOAKKYDEQAKR 351

Db 265 SGASTATNDSSTVSSGA---STATNSESSTTSSGAST-----ATN 302

Qy 352 QEBLKVSSGAGYGL-SGALLIGGIGVAVTALHRKQNPVEQTTTTTTTTTTTARTYEN 410

Db 303 SSSSTTSSGASTATNDSSTTSSGAGTATNSESSTVSGISTVTNSESSTPSSGANTATN 362

Qy 411 KPANTPAQGVVDPGSGSDTMSRRSSMASTSTFFDTS1GTVQNPYADVKT5----- 464

Db 363 SSSSTTSSGANTAT-NSSSSVSSGASTATNSESSTTSSGATATNSESSTTSSGASTAT 421

Qy 465 LHDQVPTSNSTNSYQVNGNTDSVVYTIQHPPRDITDNGARLLGNPSAGIQSTYARLAL 524

Db 422 NSDSSTTSSFASTATNNSSSSTVSSGISTVTNSESSTTSSGANTATNSSSVTS----- 474

Qy 525 SGLRHDMGULTGGSNSAANTSNNPAPG 553

Db 475 AGSGTAALTGMHTTSSHA-STAVSEAKPG 502

RESULT 10

US-10-011-795B-243

/ Sequence 243, Application US/10011795B

/ Patent No. 7012131

/ GENERAL INFORMATION:

/ APPLICANT: Baker, Kevin P.

/ APPLICANT: Botstein, David

/ APPLICANT: Desmoyers, Luc

/ APPLICANT: Baton, Dan I.

/ APPLICANT: Ferrara, Napoleone

/ APPLICANT: Fong, Sherman

/ APPLICANT: Gao, Wei-Qiang

/ APPLICANT: Goddard, Audrey

/ APPLICANT: Godowski, Paul J.

/ APPLICANT: Grimaldi, Christopher J.

/ APPLICANT: Gurney, Austin L.

/ APPLICANT: Hillan, Kenneth J.

/ APPLICANT: Pan, James

/ APPLICANT: Paoni, Nicholas F.

/ TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

/ FILE REFERENCE: P2830P1C25

/ CURRENT FILING DATE: 2001-12-07

/ PRIOR FILING DATE: 1998-09-01

/ PRIOR APPLICATION NUMBER: 60/098716

/ PRIOR FILING DATE: 1998-09-01

/ PRIOR APPLICATION NUMBER: 60/098723

/ PRIOR FILING DATE: 1998-09-01

/ PRIOR APPLICATION NUMBER: 60/098749

/ PRIOR FILING DATE: 1998-09-01

/ PRIOR APPLICATION NUMBER: 60/098750

/ PRIOR FILING DATE: 1998-09-01

/ PRIOR APPLICATION NUMBER: 60/098803

/ PRIOR FILING DATE: 1998-09-02

/ PRIOR APPLICATION NUMBER: 60/098821

/ PRIOR FILING DATE: 1998-09-02

/ PRIOR APPLICATION NUMBER: 60/098843

/ PRIOR FILING DATE: 1998-09-02

/ PRIOR APPLICATION NUMBER: 60/099536

/ PRIOR FILING DATE: 1998-09-09

/ PRIOR APPLICATION NUMBER: 60/099596

/ PRIOR FILING DATE: 1998-09-09

/ PRIOR APPLICATION NUMBER: 60/099598

/ PRIOR FILING DATE: 1998-09-09

/ Prior Application data removed - See file wrapper or PALM.

/ NUMBER OF SEQ ID NOS: 477

```
; SEQ ID NO 243
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-011-795B-243
```

```
Query Match      6.3%; Score 180; DB 3; Length 596;
Best Local Similarity 19.2%; Pred. No. 1.2e-05;
Matches 109; Conservative 90; Mismatches 272; Indels 98; Gaps 19;
```

```
QY 4 GULGNPNVNNISIPPAIPPLPSQTDGAGRGQLINSTGPGSRALFTPVNSMADSGDNRA 63
DB 13 GLLHLEAATNS-----NETSTSNANTGSSVISG-----ASTATNG 49
QY 64 SDVPGLPVPMRLAASEITLNDGFEVLHGHPLDTLNRQIGSSV-PRVETQEDGKHIAVG 122
DB 50 SSVTSSGCVSTATISGSSVTSN-GVSTIV-TNSEPHHTSSGISTATNSEFTASSGISIATN 107
QY 123 QNRGVETS--VVLSDQEVARLOSIDPEGKDFVFTGGRGAGHAMVTVASDITTEARQRL 180
DB 108 SSSSTTSSGASTATNSE-----SSTPSGASTVTNNGSSSVTSSGASTATNSESTVSSRAS 163
QY 181 ELLEPKGTGESKAGESKGVGELRESNGAENTTETQTSTSTSLRSDPKMLALGTVA 240
DB 164 TATNSESSTLSSGASTATN-SDSSTTSSGASTATNSESTTSSG-----ASTAT 211
QY 241 GLIGLAATGIYQALLTPPPDPTTDPMAASATETARDOULTEAFQNPONOKYNIDE 300
DB 212 N-----SESTVSRRASTATNSESTT--SSGASTATNSESTRTNSGAGTATNSESTTS 264
QY 301 LG-----NAIPSGVLKDDVANIIEQAKAGEAEAKQAIENNAQOKKYDEQOAKR 351
DB 265 SGASTATNDSSTVSSGA---STATNSESTTSSGAST-----ATN 302
QY 352 QEELKVVSGAGYGL-SGALLIGGIGVAVTAALHRRNOPVEQTTTTTTTTTTSARTVEN 410
DB 303 SSSSTTSSGASTATNDSSTSSGAGTATNSESTVSGISTVTNSESSTPSSGANTATN 362
QY 411 KKAANTPAQGVNDTGCSEPTMRSRSMASSTSTFPDTSIGTVONPYADVKT----- 464
DB 363 SSSSTTSSGANTAT-NSESTVSSGASTATNSESTTSSGASTATNSESTTSSGASTAT 421
QY 465 LHDQVPTNSNTSVQNMGNNTDSVYVSTIQHPPRDTTNGARLLGNPSAGISTYARLAL 524
DB 422 NSDSSTTSSSEASTATNSESTVSSGISTVTNSESSTTSSGANTATNNGSSVTS----- 474
QY 525 SGGLRHDMGGLTGGSNSAVNTSNPPAPG 553
DB 475 AGSGTAALTGMMHTTSHA-STAVSEAKPG 502
```

RESULT 11

```
US-10-015-386A-243
; Sequence 243, Application US/10015386A
; Patent No. 7022498
```

GENERAL INFORMATION:

```
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan I.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Pan, James
; APPLICANT: Paoli, Nicholas F.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2830PIC55
```

```
; CURRENT APPLICATION NUMBER: US/10/015,386A
; CURRENT FILING DATE: 2001-12-12
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 477
; SEQ ID NO 243
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-015-386A-243
```

```
Query Match      6.3%; Score 180; DB 3; Length 596;
Best Local Similarity 19.2%; Pred. No. 1.2e-05;
Matches 109; Conservative 90; Mismatches 272; Indels 98; Gaps 19;
```

```
QY 4 GULGNPNVNNISIPPAIPPLPSQTDGAGRGQLINSTGPGSRALFTPVNSMADSGDNRA 63
DB 13 GLLHLEAATNS-----NETSTSNANTGSSVISG-----ASTATNG 49
QY 64 SDVPGLPVPMRLAASEITLNDGFEVLHGHPLDTLNRQIGSSV-PRVETQEDGKHIAVG 122
DB 50 SSVTSSGCVSTATISGSSVTSN-GVSTIV-TNSEPHHTSSGISTATNSEFTASSGISIATN 107
QY 123 QNRGVETS--VVLSDQEVARLOSIDPEGKDFVFTGGRGAGHAMVTVASDITTEARQRL 180
DB 108 SSSSTTSSGASTATNSE-----SSTPSGASTVTNNGSSSVTSSGASTATNSESTVSSRAS 163
QY 181 ELLEPKGTGESKAGESKGVGELRESNGAENTTETQTSTSTSLRSDPKMLALGTVA 240
DB 164 TATNSESSTLSSGASTATN-SDSSTTSSGASTATNSESTTSSG-----ASTAT 211
QY 241 GLIGLAATGIYQALLTPPPDPTTDPMAASATETARDOULTEAFQNPONOKYNIDE 300
DB 212 N-----SESTVSRRASTATNSESTT--SSGASTATNSESTRTNSGAGTATNSESTTS 264
QY 301 LG-----NAIPSGVLKDDVANIIEQAKAGEAEAKQAIENNAQOKKYDEQOAKR 351
DB 265 SGASTATNDSSTVSSGA---STATNSESTTSSGAST-----ATN 302
QY 352 QEELKVVSGAGYGL-SGALLIGGIGVAVTAALHRRNOPVEQTTTTTTTTTTSARTVEN 410
DB 303 SSSSTTSSGASTATNDSSTSSGAGTATNSESTVSGISTVTNSESSTPSSGANTATN 362
QY 411 KKAANTPAQGVNDTGCSEPTMRSRSMASSTSTFPDTSIGTVONPYADVKT----- 464
DB 363 SSSSTTSSGANTAT-NSESTVSSGASTATNSESTTSSGASTATNSESTTSSGASTAT 421
QY 465 LHDQVPTNSNTSVQNMGNNTDSVYVSTIQHPPRDTTNGARLLGNPSAGISTYARLAL 524
DB 422 NSDSSTTSSSEASTATNSESTVSSGISTVTNSESSTTSSGANTATNNGSSVTS----- 474
QY 525 SGGLRHDMGGLTGGSNSAVNTSNPPAPG 553
DB 475 AGSGTAALTGMMHTTSHA-STAVSEAKPG 502
```

RESULT 12

```
US-10-012-121A-243
; Sequence 243, Application US/10012121A
; Patent No. 7022817
```

GENERAL INFORMATION:

```
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan I.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Pan, James
```

```

; APPLICANT: Paoni, Nicholas F.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2830P1C20
; CURRENT APPLICATION NUMBER: US/10/012,121A
; PRIOR APPLICATION REMOVED - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 477
; SEQ ID NO 243
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-012-121A-243

Query Match      6.3%; Score 180; DB 3; Length 596;
Best Local Similarity 19.2%; Pred. No. 1,2e-05;
Matches 109; Conservative 90; Mismatches 272; Indels 98; Gaps 19;

QY      4 GNLGNPNVNNISIPAPLPISQTDGAGRGQLINSTPLGSRALFTPVNMSADSDNRA 63
DB      13 GILLHLEAATNS-----NETSTISANTGSSVSISSG-----ASTATNSG 49
QY      64 SDVPLPVNPKMLAASEITLNDGFEVLHDGPLDILNRQIGSSV-FRVETQEDGHIANG 122
DB      50 SSVTSSGVSTATISGSSVTSN-GVSIIV-TNSEPHITSSGISTATNSEFSTAGSISIAIN 107
QY      123 QNNGVETS--VLSQOEVARLOSIDPEGKDKVFTFGGAGAGHAMVTVASDITTEARQRL 180
DB      108 SSSSTTSSGASTATNSE--SSTPSSGASTVYTNSSGSSVTSAGASTATNSESSTVSSRS 163
QY      181 ELLEPKGTGESSKAGSGKGVGELRESNGAENTETOTSTSSLSRDPKMLALGTIVAT 240
DB      164 TATNSESSTLSGASTATN-SDSSTTSSGASTATNSESSTTSSG-----ASTAT 211
QY      241 GLIGLAATGIVQALATPEPDSPITTPDPAASATETATRDQLTKAFQNPQNKVNIDE 300
DB      212 N-----SESSTVSSRASTATNSESSTT--SGASTATNSESRTTSGAGTATNSESSTTS 264
QY      301 LG-----NAISGVAKDVIIVANIEQAAAGEBAKQQAILENNAQOKKYDDQAKR 351
DB      265 SGASTATNDSSTVSSGA---STATNSESSTTSSGAST-----ATN 302
QY      352 QRELKVSAGAGYGL-SGAILGGIGVAVTALHRAKQVFEQTTTTTTTTTARTVEN 410
DB      303 SSSSTTSSGASTATNDSSTTSSGAGTATNSESSTVSSISIVTVNSESSTPSSGANTAIN 362
QY      411 KPANNTPAQGNVDTPGSEDTMESRRSSMASTSTFEPTISSIGTVONPVADVTS----- 464
DB      363 SSSSTTSSGANTAT-NSESSTVSSGASTATNSESSTTSSGASTATNSESSTTSSGASTAT 421
QY      465 LHDQVPTNSNTSVQNMAMTDSVYVSTIQHPRTDTNGALLGNPSAGIOSTARLAL 524
DB      422 NSDSSTTSSGASTATNSESSTVSSGISTVTSNSESSTTSSGANTATNSESSTVSS 474
QY      525 SGGLRHDMGGLTGGSNSAVTNNPPAPG 553
DB      475 AGSGTALTGMTTSHSA-STAVSEAKPG 502

RESULT 13
US-10-006-485A-243
; Sequence 243, Application US/10006485A
; Patent No. 7026448
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Eaton, Dan I.
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
```

```

; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gunney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2830P1C9
; CURRENT APPLICATION NUMBER: US/10/006,485A
; CURRENT FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: 60/098716
; PRIOR FILING DATE: 1998-09-01
; PRIOR APPLICATION NUMBER: 60/098723
; PRIOR FILING DATE: 1998-09-01
; PRIOR APPLICATION NUMBER: 60/098749
; PRIOR FILING DATE: 1998-09-01
; PRIOR APPLICATION NUMBER: 60/098750
; PRIOR FILING DATE: 1998-09-01
; PRIOR APPLICATION NUMBER: 60/098803
; PRIOR FILING DATE: 1998-09-02
; PRIOR APPLICATION NUMBER: 60/098821
; PRIOR FILING DATE: 1998-09-02
; PRIOR APPLICATION NUMBER: 60/098843
; PRIOR FILING DATE: 1998-09-02
; PRIOR APPLICATION NUMBER: 60/099536
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: 60/099596
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: 60/099598
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: 60/099602
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: 60/099642
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: 60/099741
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099754
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099763
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099792
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099808
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099812
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099815
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099816
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/100385
; PRIOR FILING DATE: 1998-09-15
; PRIOR APPLICATION NUMBER: 60/100388
; PRIOR FILING DATE: 1998-09-15
; PRIOR APPLICATION NUMBER: 60/100390
; PRIOR FILING DATE: 1998-09-15
; PRIOR APPLICATION NUMBER: 60/100584
; PRIOR FILING DATE: 1998-09-16
; PRIOR APPLICATION NUMBER: 60/100627
; PRIOR FILING DATE: 1998-09-16
; PRIOR APPLICATION NUMBER: 60/100661
; PRIOR FILING DATE: 1998-09-16
; PRIOR APPLICATION NUMBER: 60/100662
; PRIOR FILING DATE: 1998-09-16
; PRIOR APPLICATION NUMBER: 60/100664
; PRIOR FILING DATE: 1998-09-16
; PRIOR APPLICATION NUMBER: 60/100683
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: 60/100684
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: 60/100710
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: 60/100711
```


PRIOR FILING DATE: 1998-09-17
PRIOR APPLICATION NUMBER: 60/100848
PRIOR FILING DATE: 1998-09-18
PRIOR APPLICATION NUMBER: 60/100849
PRIOR FILING DATE: 1998-09-18
PRIOR APPLICATION NUMBER: 60/100919
PRIOR FILING DATE: 1998-09-17
PRIOR APPLICATION NUMBER: 60/100930
PRIOR FILING DATE: 1998-09-17
PRIOR APPLICATION NUMBER: 60/101014
PRIOR FILING DATE: 1998-09-18
PRIOR APPLICATION NUMBER: 60/101068
PRIOR FILING DATE: 1998-09-18
PRIOR APPLICATION NUMBER: 60/101071
PRIOR FILING DATE: 1998-09-18
PRIOR APPLICATION NUMBER: 60/101279
PRIOR FILING DATE: 1998-09-22
PRIOR APPLICATION NUMBER: 60/101471
PRIOR FILING DATE: 1998-09-23
PRIOR APPLICATION NUMBER: 60/101472
PRIOR FILING DATE: 1998-09-23
PRIOR APPLICATION NUMBER: 60/101474
PRIOR FILING DATE: 1998-09-23
PRIOR APPLICATION NUMBER: 60/101475
PRIOR FILING DATE: 1998-09-23
PRIOR APPLICATION NUMBER: 60/101476
PRIOR FILING DATE: 1998-09-23
PRIOR APPLICATION NUMBER: 60/101477
PRIOR FILING DATE: 1998-09-23
PRIOR APPLICATION NUMBER: 60/101479
PRIOR FILING DATE: 1998-09-23
PRIOR APPLICATION NUMBER: 60/101738
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/101741
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/101743
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/101915
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/101916
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/102207
PRIOR FILING DATE: 1998-09-29
PRIOR APPLICATION NUMBER: 60/102240
PRIOR FILING DATE: 1998-09-29
PRIOR APPLICATION NUMBER: 60/102307
PRIOR FILING DATE: 1998-09-29
PRIOR APPLICATION NUMBER: 60/102330
PRIOR FILING DATE: 1998-09-29
PRIOR APPLICATION NUMBER: 60/102331
PRIOR FILING DATE: 1998-09-29
PRIOR APPLICATION NUMBER: 60/102484
PRIOR FILING DATE: 1998-09-30
PRIOR APPLICATION NUMBER: 60/102487
PRIOR FILING DATE: 1998-09-30
PRIOR APPLICATION NUMBER: 60/102570
PRIOR FILING DATE: 1998-09-30
PRIOR APPLICATION NUMBER: 60/102571
PRIOR FILING DATE: 1998-09-30
PRIOR APPLICATION NUMBER: 60/102684
PRIOR FILING DATE: 1998-10-01
PRIOR APPLICATION NUMBER: 60/102687
PRIOR FILING DATE: 1998-10-01
PRIOR APPLICATION NUMBER: 60/102965
PRIOR FILING DATE: 1998-10-02
PRIOR APPLICATION NUMBER: 60/103258
PRIOR FILING DATE: 1998-10-06
PRIOR APPLICATION NUMBER: 60/103314
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: 60/103315
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: 60/103328
PRIOR FILING DATE: 1998-10-07

PRIOR APPLICATION NUMBER: 60/103395
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: 60/103396
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: 60/103401
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: 60/103449
PRIOR FILING DATE: 1998-10-06
PRIOR APPLICATION NUMBER: 60/103633
PRIOR FILING DATE: 1998-10-08
PRIOR APPLICATION NUMBER: 60/103678
PRIOR FILING DATE: 1998-10-08
PRIOR APPLICATION NUMBER: 60/103679
PRIOR FILING DATE: 1998-10-08
PRIOR APPLICATION NUMBER: 60/103711
PRIOR FILING DATE: 1998-10-08
PRIOR APPLICATION NUMBER: 60/104257
PRIOR FILING DATE: 1998-10-14
PRIOR APPLICATION NUMBER: 60/104987
PRIOR FILING DATE: 1998-10-20
PRIOR APPLICATION NUMBER: 60/105000
PRIOR FILING DATE: 1998-10-20
PRIOR APPLICATION NUMBER: 60/105002
PRIOR FILING DATE: 1998-10-20
PRIOR APPLICATION NUMBER: 60/105104
PRIOR FILING DATE: 1998-10-21
PRIOR APPLICATION NUMBER: 60/105169
PRIOR FILING DATE: 1998-10-22
PRIOR APPLICATION NUMBER: 60/105266
PRIOR FILING DATE: 1998-10-22
PRIOR APPLICATION NUMBER: 60/105693
PRIOR FILING DATE: 1998-10-26
PRIOR APPLICATION NUMBER: 60/105694
PRIOR FILING DATE: 1998-10-26
PRIOR APPLICATION NUMBER: 60/105807
PRIOR FILING DATE: 1998-10-27
PRIOR APPLICATION NUMBER: 60/105881
PRIOR FILING DATE: 1998-10-27
PRIOR APPLICATION NUMBER: 60/105882
PRIOR FILING DATE: 1998-10-27
PRIOR APPLICATION NUMBER: 60/106023
PRIOR FILING DATE: 1998-10-28

Query Match 6.3%; Score 180; DB 3; Length 596;
Best Local Similarity 19.2%; Pred. No. 1.2e-05;
Matches 109; Conservative 90; Mismatches 272; Indels 98; Gaps 19;

4 GNLGNPNVNNISPPAPLPISQTDGAGRGQJLNSGPIGSRALFTFPVNSMADSGDNRA 63
13 GLLHLHAATNS-----NETSTANTGSSVTSIG-----ASTATNSG 49
64 SDVPGLPVPMRLAASEITLNDGFEVLHDGPIDLTLNROIQSSV-PRVETQEDGKHIAVG 122
50 SSVTSSGCVSTATISGSVTSN-CVSI-V-TNSEFHTSSSISTRTATNSBFTASGISTATN 107
123 QNRNGVTS--VVLSDQEVARLQSIDEGKDKVFTEGRCGAGAMVTVAASDITEARQRI 180
108 SSSSTSSGASTATNSE-----SSTPSSGASTVNSGSSVTSAGASTATNSSESTVSSRAS 163
181 ELLEPKGTESGAGSKVGVGLRENSGAEVTEIQTSTSSLSRSDPKMLALGTVA 240
164 TATNSSESTLSSGASTATN-SDSSTSSGASTATNSSESTSSG-----ASTAT 211
241 GILGLAATGIVQALATPEPDSPTTTPDPAASATETATRDOLTKAFOPNQKVNIDE 300
212 N-----SESSTVSSRASTANSESTT---SSGASTATNSSESTTSGAGTATNSSESTTS 264
301 LG-----NAIPSGVILKDDVAVNIEEQAVAGEBAKQQAENNAQOKKYDEQQAAR 351
265 SGASTATNSDSSTVSSGA---STATNSESTTSSGAST-----ATN 302
352 QELKVSAGAGVGL-SGAILLOGGIGVAVTALHKNQVDEQTTTTTTTTTTSARTVFN 410

Db 303 SSSSTSSGASTATNSDSTSSGAGTATNSSESTVSSGISTVTNSSESTPSSGANTATN 362
Qy 411 KPANNTAQANDTGCSEDDTMSRRSSMASTSTFPDSSICTVONPADVITS----- 464
Db 363 SSSSTSSGANTAT-NSBSTVSSGASTATNSSESTSSGVTATNSSESTSSGASTAT 421
Qy 465 LHDSCVPTNSNTSVONKMGNTDSVVYSTIOHPRTTONGARLLGNPSAGIOSTYARLAL 524
Db 422 NSDSTSSSEASTANSSSESTVSSGISTVTNSSESTSSGANTATNSGSVTS----- 474
Qy 525 SGGRLHDMGCLTGGSNSAVTNSNPPAPG 553
Db 475 AGSGTALTGMTTSHSA-STAVSEAKPG 502

RESULT 14
US-10-006-746A-243
Sequence 243, Application US/10006746A
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan I.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE OF INVENTION: Acids Encoding the Same
FILE REFERENCE: P2830PLC5
CURRENT APPLICATION NUMBER: US/10/006,746A
CURRENT FILING DATE: 2001-12-06
PRIOR APPLICATION NUMBER: 60/098716
PRIOR FILING DATE: 1998-09-01
PRIOR APPLICATION NUMBER: 60/098723
PRIOR FILING DATE: 1998-09-01
PRIOR APPLICATION NUMBER: 60/098749
PRIOR FILING DATE: 1998-09-01
PRIOR APPLICATION NUMBER: 60/098750
PRIOR FILING DATE: 1998-09-01
PRIOR APPLICATION NUMBER: 60/098803
PRIOR FILING DATE: 1998-09-02
PRIOR APPLICATION NUMBER: 60/098821
PRIOR FILING DATE: 1998-09-02
PRIOR APPLICATION NUMBER: 60/098843
PRIOR FILING DATE: 1998-09-02
PRIOR APPLICATION NUMBER: 60/099536
PRIOR FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: 60/099596
PRIOR FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: 60/099598
PRIOR FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: 60/099602
PRIOR FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: 60/099642
PRIOR FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: 60/099741
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099754
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099763
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099792
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099808
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/102240

PRIOR APPLICATION NUMBER: 60/099812
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099815
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099816
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/100385
PRIOR FILING DATE: 1998-09-15
PRIOR APPLICATION NUMBER: 60/100388
PRIOR FILING DATE: 1998-09-15
PRIOR APPLICATION NUMBER: 60/100390
PRIOR FILING DATE: 1998-09-15
PRIOR APPLICATION NUMBER: 60/100584
PRIOR FILING DATE: 1998-09-16
PRIOR APPLICATION NUMBER: 60/100627
PRIOR FILING DATE: 1998-09-16
PRIOR APPLICATION NUMBER: 60/100661
PRIOR FILING DATE: 1998-09-16
PRIOR APPLICATION NUMBER: 60/100662
PRIOR FILING DATE: 1998-09-16
PRIOR APPLICATION NUMBER: 60/100664
PRIOR FILING DATE: 1998-09-16
PRIOR APPLICATION NUMBER: 60/100683
PRIOR FILING DATE: 1998-09-17
PRIOR APPLICATION NUMBER: 60/100684
PRIOR FILING DATE: 1998-09-17
PRIOR APPLICATION NUMBER: 60/100710
PRIOR FILING DATE: 1998-09-17
PRIOR APPLICATION NUMBER: 60/100711
PRIOR FILING DATE: 1998-09-17
PRIOR APPLICATION NUMBER: 60/100848
PRIOR FILING DATE: 1998-09-18
PRIOR APPLICATION NUMBER: 60/100849
PRIOR FILING DATE: 1998-09-18
PRIOR APPLICATION NUMBER: 60/100919
PRIOR FILING DATE: 1998-09-17
PRIOR APPLICATION NUMBER: 60/100930
PRIOR FILING DATE: 1998-09-17
PRIOR APPLICATION NUMBER: 60/101014
PRIOR FILING DATE: 1998-09-18
PRIOR APPLICATION NUMBER: 60/101068
PRIOR FILING DATE: 1998-09-18
PRIOR APPLICATION NUMBER: 60/101071
PRIOR FILING DATE: 1998-09-18
PRIOR APPLICATION NUMBER: 60/101279
PRIOR FILING DATE: 1998-09-22
PRIOR APPLICATION NUMBER: 60/101471
PRIOR FILING DATE: 1998-09-23
PRIOR APPLICATION NUMBER: 60/101472
PRIOR FILING DATE: 1998-09-23
PRIOR APPLICATION NUMBER: 60/101474
PRIOR FILING DATE: 1998-09-23
PRIOR APPLICATION NUMBER: 60/101475
PRIOR FILING DATE: 1998-09-23
PRIOR APPLICATION NUMBER: 60/101476
PRIOR FILING DATE: 1998-09-23
PRIOR APPLICATION NUMBER: 60/101477
PRIOR FILING DATE: 1998-09-23
PRIOR APPLICATION NUMBER: 60/101479
PRIOR FILING DATE: 1998-09-23
PRIOR APPLICATION NUMBER: 60/101738
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/101741
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/101743
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/101915
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/101916
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/102207
PRIOR FILING DATE: 1998-09-29
PRIOR APPLICATION NUMBER: 60/102240

Query Match 6.3% Score 180; DB 3; Length 596;

PRIOR FILING DATE: 1998-09-29
PRIOR APPLICATION NUMBER: 60/102307
PRIOR FILING DATE: 1998-09-29
PRIOR APPLICATION NUMBER: 60/102330
PRIOR FILING DATE: 1998-09-29
PRIOR APPLICATION NUMBER: 60/102331
PRIOR FILING DATE: 1998-09-29
PRIOR APPLICATION NUMBER: 60/102484
PRIOR FILING DATE: 1998-09-30
PRIOR APPLICATION NUMBER: 60/102487
PRIOR FILING DATE: 1998-09-30
PRIOR APPLICATION NUMBER: 60/102570
PRIOR FILING DATE: 1998-09-30
PRIOR APPLICATION NUMBER: 60/102571
PRIOR FILING DATE: 1998-09-30
PRIOR APPLICATION NUMBER: 60/102684
PRIOR FILING DATE: 1998-10-01
PRIOR APPLICATION NUMBER: 60/102687
PRIOR FILING DATE: 1998-10-01
PRIOR APPLICATION NUMBER: 60/102965
PRIOR FILING DATE: 1998-10-02
PRIOR APPLICATION NUMBER: 60/103258
PRIOR FILING DATE: 1998-10-06
PRIOR APPLICATION NUMBER: 60/103314
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: 60/103315
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: 60/103328
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: 60/103395
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: 60/103396
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: 60/103401
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: 60/103449
PRIOR FILING DATE: 1998-10-06
PRIOR APPLICATION NUMBER: 60/103633
PRIOR FILING DATE: 1998-10-08
PRIOR APPLICATION NUMBER: 60/103678
PRIOR FILING DATE: 1998-10-08
PRIOR APPLICATION NUMBER: 60/103679
PRIOR FILING DATE: 1998-10-08
PRIOR APPLICATION NUMBER: 60/103711
PRIOR FILING DATE: 1998-10-08
PRIOR APPLICATION NUMBER: 60/104257
PRIOR FILING DATE: 1998-10-14
PRIOR APPLICATION NUMBER: 60/104987
PRIOR FILING DATE: 1998-10-20
PRIOR APPLICATION NUMBER: 60/105000
PRIOR FILING DATE: 1998-10-20
PRIOR APPLICATION NUMBER: 60/105002
PRIOR FILING DATE: 1998-10-20
PRIOR APPLICATION NUMBER: 60/105104
PRIOR FILING DATE: 1998-10-21
PRIOR APPLICATION NUMBER: 60/105169
PRIOR FILING DATE: 1998-10-22
PRIOR APPLICATION NUMBER: 60/105266
PRIOR FILING DATE: 1998-10-22
PRIOR APPLICATION NUMBER: 60/105693
PRIOR FILING DATE: 1998-10-26
PRIOR APPLICATION NUMBER: 60/105694
PRIOR FILING DATE: 1998-10-26
PRIOR APPLICATION NUMBER: 60/105697
PRIOR FILING DATE: 1998-10-27
PRIOR APPLICATION NUMBER: 60/105881
PRIOR FILING DATE: 1998-10-27
PRIOR APPLICATION NUMBER: 60/105882
PRIOR FILING DATE: 1998-10-27
PRIOR APPLICATION NUMBER: 60/106023
PRIOR FILING DATE: 1998-10-28

Best Local Similarity 19.2%; Pred. No. 1,2e-05;
Matches 109; Conservative 90; Mismatches 272; Indels 98; Gaps 19;

QY 4 GNIGHNPNVNSIPPAPPLPSQTDGAGRGQLINSTGPGSLRALFPVPRNSMADSGDNRA 63
13 GLLHLEAATNS-----NETSTANTSSVSSG-----ASTATNSG 49
QY 64 SDVPGLPVPMRLAAEITLNDGFVLHDHGPJDTLNRQSSV-FRVEIQEGKHAVG 122
50 SSVTSSGVSTATSSSVTSN-GVSI-V-TNSEPHSTSGISTATNSEFSTASSGISIATN 107
QY 123 ORNGVETS--VVLSDPEVARLOSIDPEKDKFVGTGRGAGAHMTVVASDITFARQIL 180
108 SESSSTSSGASTATNE---STPSSGASTVNSSSVSSGASTATNSESSVSSRAS 163
QY 181 ELLEPKGTSGSKAGESKGVGBLRESNSGAENTETOTSTSTSLRSDPKMLALGTVA 240
164 TATNSSTSLSSGASTATN-SDSSTSSGASTATNSESSSTSG-----ASTAT 211
QY 241 GLIGLAATGVQALATTPPDSTTTDPDAASATETARDQUTKEAFQNPQKNIDE 300
212 N---SESTVSSRASTATNSESTT---SSGASTATNSESTTNGAGTATNSESTTS 264
QY 301 LG-----NAIPSGVLKDPVAVNIEBOAKAAGEAKQAENNAQAQKXDEQAKR 351
265 SGASTATNSDSSVSSGA---STATNSESTSSGAST-----ATN 302
QY 352 QEELKVSAGAGYGL-SCALILGGIGVAVTAAHNRKQVEQOTTTTTTTTSARTVEN 410
303 SSSSTSSGASTATNSDSSSTSSGAGTATNSESTVSSGISTVTSSESTPSSGANTATN 362
QY 411 KPAANTPAQGVNDTPGSEPTMESRRSSMASTSTPDTSTIGTVQNPVADKTS----- 464
363 SSSSTSSGANTAT-NSSESTVSSGASTATNSESTSSGASTATNSESTSSGASTAT 421
QY 465 LHDQVPTNSNMTSVQNMONTDSVYSTIQHPRODTDNGARLGNPSAGICQTVARLAL 524
422 NSDSTSSSEASTATNSESTVSSGISTVTSSESTSSGANTATNSGSSVTS----- 474
QY 525 SGGLRHDMGLTGCNSAVNTSNPPAPG 553
475 AGSGTALTGWHHTTSHA-STAVSEAKPG 502
Db

RESULT 15
US-10-012-752A-243
; Sequence 243, Application US/10012752A
; Patent No. 7026455
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Baton, Dan I.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2830PIC24
; CURRENT APPLICATION NUMBER: US/10/012,752A
; CURRENT FILING DATE: 2002-06-25
; Prior application removed - See File Wrapper or Palm
; SEQ ID NO 243
; LENGTH: 596
; TYPE: PRT

ORGANISM: Homo sapiens
US-10-012-752A-243

Query Match 6.3%; Score 180; DB 3; Length 596;
Best Local Similarity 19.2%; Pred. No. 1.2e-05;
Matches 109; Conservative 90; Mismatches 272; Indels 98; Gaps 19;

```
OY 4 GNIGNPNVNNISIPAPPLPSQTDGAGRGQLINSTGLSGRALFTPNRNSMADSGDNRA 63
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 13 GILHLLEAATNS-----NETSTANTGSSVSISSG-----ASTATNSG 49

OY 64 SDVPELPVNPMLAASEITLNDGFEVLHDHGELDLTINRQIGSSV-FRVETQEDGKHIAVG 122
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 50 SSVTSSGVTATISGSSVTSN-GVAIV-TNSEFHTSSGISTATNSFSTASSGISTATN 107

OY 123 QRNGVETS--VLSQDEYARLOSIDPEGDKFVFTGGRGAGHAMVTVASDITEARQRL 180
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 108 SSSSTTSSGASTATNSE-----STPSSGASTVTVNCGSSVTSAGASTATNSSTVSSRAS 163

OY 181 ELLEPKGGESEKAGESKGVGELRESNGAENTTETOTSTSTSLRSDPKMLALGTAT 240
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 164 TATNSSESTLSSGASTATN-SDSSTTSSGASTATNSSESTTSSG-----ASTAT 211

OY 241 GLIGLAATGIYQALATPEPDSPITTPDPAASATEATARDQLYKEAFQNPNOKVNIIDE 300
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 212 N-----SESTVSSRASTATNSSESTT---SSGASTATNSSESTTNGAGTATNSSESTTS 264

OY 301 LG-----NAIPSGVLKDDVANIIEBQAKAGEBAKQQA IENNAQAKKYDEQAKR 351
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 265 SGASTATNSDSSTVSSGA---STATNSSESTTSSGAST-----ATN 302

OY 352 OEELKVVSSGAGYGL--SGALLGGGIVAVTALHKNQVPEQTTTTTTTTTSARTVEN 410
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 303 SSSSTTSSGASTATNSDSSTTSSGAGTATNSSESTVSSGISTVTNSSESTPSSGANTATN 362

OY 411 KPAANTPAQGVNTPGSEDTMESRRSMASTSTFFDTSIGTVQNPYADVKT----- 464
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 363 SSSSTTSSGANTAT-NSSESTVSSGASTATNSSESTTSSGVTATNSSESTTSSGASTAT 421

OY 465 LHDQVPTNSNTSVQNMGNATDSVVYSTIQHPPTDITNGARLLGNPSAGIOSYARLAL 524
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 422 NSDSSTTSSASTATNSSESTVSSGISTVTNSSESTTSSGANTATNSGSSVTS----- 474

OY 525 SGGLRHDMDGLTGGSNSAVNTSNPPAPG 553
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 475 AGSGTALTGMHTTSHA-STAVSEAKPG 502
```

Search completed: August 1, 2006, 22:14:59
Job time : 53 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: August 1, 2006, 22:25:31 ; Search time 184 Seconds
(without alignments)
1404.749 Million cell updates/sec

Title: US-09-189-415d-11

Perfect score: 2840

Sequence: 1 MFIGNLGNHPNNVNSIPAP.....SNGAVNTSNNPAPGSHRTV 558

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2097797 seqs, 463214858 residues

Total number of hits satisfying chosen parameters: 2097797

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA Main:*

- 1: /EMC_Celerra_SIDS3/ptodata/2/pubppaa/US07_PUBCOMB.pep:*
- 2: /EMC_Celerra_SIDS3/ptodata/2/pubppaa/US08_PUBCOMB.pep:*
- 3: /EMC_Celerra_SIDS3/ptodata/2/pubppaa/US09_PUBCOMB.pep:*
- 4: /EMC_Celerra_SIDS3/ptodata/2/pubppaa/US10_PUBCOMB.pep:*
- 5: /EMC_Celerra_SIDS3/ptodata/2/pubppaa/US10B_PUBCOMB.pep:*
- 6: /EMC_Celerra_SIDS3/ptodata/2/pubppaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2840	100.0	558	US-11-134-563-12	Sequence 12, Appl
2	2840	100.0	558	US-11-052-554A-20	Sequence 20, Appl
3	187	6.6	1300	US-11-097-143-31017	Sequence 31017, A
4	182.5	6.4	2586	US-11-097-143-27426	Sequence 27426, A
5	180	6.3	596	US-09-946-374-243	Sequence 243, App
6	180	6.3	596	US-10-006-867-100	Sequence 100, App
7	180	6.3	596	US-10-052-586-310	Sequence 310, App
8	180	6.3	596	US-10-063-547-100	Sequence 100, App
9	180	6.3	596	US-10-063-551-100	Sequence 100, App
10	180	6.3	596	US-10-174-590-310	Sequence 310, App
11	180	6.3	596	US-10-176-758-310	Sequence 310, App
12	180	6.3	596	US-10-175-737-310	Sequence 310, App
13	180	6.3	596	US-10-063-616-100	Sequence 100, App
14	180	6.3	596	US-10-174-581-310	Sequence 310, App
15	180	6.3	596	US-10-176-483-310	Sequence 310, App
16	180	6.3	596	US-10-176-749-310	Sequence 310, App
17	180	6.3	596	US-10-176-914-310	Sequence 310, App
18	180	6.3	596	US-10-176-915-310	Sequence 310, App
19	180	6.3	596	US-10-063-569-100	Sequence 100, App
20	180	6.3	596	US-10-063-513-100	Sequence 100, App
21	180	6.3	596	US-10-063-515-100	Sequence 100, App
22	180	6.3	596	US-10-063-512-100	Sequence 100, App
23	180	6.3	596	US-10-173-706-310	Sequence 310, App
24	180	6.3	596	US-10-175-738-310	Sequence 310, App
25	180	6.3	596	US-10-175-752-310	Sequence 310, App
26	180	6.3	596	US-10-176-482-310	Sequence 310, App
27	180	6.3	596	US-10-176-757-310	Sequence 310, App

ALIGNMENTS

28	180	6.3	596	4	US-10-176-913-310	Sequence 310, App
29	180	6.3	596	4	US-10-180-552-310	Sequence 310, App
30	180	6.3	596	4	US-10-180-557-310	Sequence 310, App
31	180	6.3	596	4	US-10-063-502-100	Sequence 100, App
32	180	6.3	596	4	US-10-173-700-310	Sequence 310, App
33	180	6.3	596	4	US-10-174-572-310	Sequence 310, App
34	180	6.3	596	4	US-10-174-579-310	Sequence 310, App
35	180	6.3	596	4	US-10-174-582-310	Sequence 310, App
36	180	6.3	596	4	US-10-174-588-310	Sequence 310, App
37	180	6.3	596	4	US-10-175-739-310	Sequence 310, App
38	180	6.3	596	4	US-10-175-740-310	Sequence 310, App
39	180	6.3	596	4	US-10-175-743-310	Sequence 310, App
40	180	6.3	596	4	US-10-176-488-310	Sequence 310, App
41	180	6.3	596	4	US-10-176-492-310	Sequence 310, App
42	180	6.3	596	4	US-10-176-747-310	Sequence 310, App
43	180	6.3	596	4	US-10-176-750-310	Sequence 310, App
44	180	6.3	596	4	US-10-176-985-310	Sequence 310, App
45	180	6.3	596	4	US-10-176-987-310	Sequence 310, App

RESULT 1
US-11-134-563-12
; Sequence 12, Application US/11134563
; Publication No. US20050287569A1
; GENERAL INFORMATION:
; APPLICANT: Leong, John M.
; TITLE OF INVENTION: ESPEU NUCLEIC ACIDS AND PROTEINS AND
; FILE REFERENCE: 07917-280001
; CURRENT APPLICATION NUMBER: US/11/134,563
; PRIOR FILING DATE: 2005-05-20
; PRIOR APPLICATION NUMBER: US 60/573,600
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 558
; TYPE: PR1
; ORGANISM: Escherichia coli
US-11-134-563-12

Query Match 100.0%; Score 2840; DB 6; Length 558;
Best Local Similarity 100.0%; Pred. No. 3.6e-177;
Matches 558; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MFIGNLGNHPNNVNSIPAPPLPSQTDGAGRGQLNSTGPGSRALFTPVNSMADSGD	60
DB	1	MFIGNLGNHPNNVNSIPAPPLPSQTDGAGRGQLNSTGPGSRALFTPVNSMADSGD	60
QY	61	NRASDPVGLPVNPMRLAASEITLNDGFEVLHGHGPDLTNRQIGSSVFVEVQEDKHIA	120
DB	61	NRASDPVGLPVNPMRLAASEITLNDGFEVLHGHGPDLTNRQIGSSVFVEVQEDKHIA	120
QY	121	VCORNGVETSVVLSDOEYARLOSIDPEGKDFEFTGRCGAGHAMVTVASDITTEARQIL	180
DB	121	VCORNGVETSVVLSDOEYARLOSIDPEGKDFEFTGRCGAGHAMVTVASDITTEARQIL	180
QY	181	ELLEPKGTESKAGAGSKVGLRESNSGAENTTETOTSTSSLSRSDPKLWALGTVAT	240
DB	181	ELLEPKGTESKAGAGSKVGLRESNSGAENTTETOTSTSSLSRSDPKLWALGTVAT	240
QY	241	GLIGLAANGIYVALALTPEDSPPTTDPDAASATETARRDQTYKAFQNPONOKNIDE	300
DB	241	GLIGLAANGIYVALALTPEDSPPTTDPDAASATETARRDQTYKAFQNPONOKNIDE	300
QY	301	LGNAIPSGVLKQDVVANIEBQAKAAGEAKQQAIEENNAQAKKYDEBQAKROBELKVSSG	360
DB	301	LGNAIPSGVLKQDVVANIEBQAKAAGEAKQQAIEENNAQAKKYDEBQAKROBELKVSSG	360

Qy	AGGAGSGLILGGIGIVAVTAALHRRNQVEOTTTTTTTTTTSARTYENKPAANNPAAG	420
Db	AGGAGSGLILGGIGIVAVTAALHRRNQVEOTTTTTTTTTTSARTYENKPAANNPAAG	420
Qy	NVDTPGSEDTEMRSSMASTSTPEFDISLITVONPADYKTSLHDSQVPTSNSNTSVQ	480
Db	NVDTPGSEDTEMRSSMASTSTPEFDISLITVONPADYKTSLHDSQVPTSNSNTSVQ	480
Qy	NMGNTDSVYVSTIQHPPRDITDNGARLLGNPSAGIQSTYARIALSGSLRHDGGLTGGSN	540
Db	NMGNTDSVYVSTIQHPPRDITDNGARLLGNPSAGIQSTYARIALSGSLRHDGGLTGGSN	540
Qy	SAVNTSNPPAPGSHRFV	558
Db	SAVNTSNPPAPGSHRFV	558

```

RESULT 2
US-11-052-554A--20
; Sequence 20, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; TITLE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; PRIOR FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 20
; LENGTH: 558
; TYPE: PRT
; ORGANISM: Escherichia coli 0157:H7
US-11-052-554A--20

```

Query Match	100.0%	Score 2840	DB 6	Length 558
Best Local Similarity	100.0%	Pred. No. 3,66-177		
Matches 558	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY	1	MPIGNLGNPNVNNISIPAPPLPSQTDGAGRGOLINS7GPLGRALLFTPVNSMADSGD	60	
Db	1	MPIGNLGNPNVNNISIPAPPLPSQTDGAGRGQLINS7GPLGRALLFTPVNSMADSGD	60	
QY	61	NRASVPEPLPNVPMPLAASEITLNGFEVLJHGPDLTINRQIGSSVFPEVQEDSKHA	120	
Db	61	NRASVPEPLPNVPMPLAASEITLNGFEVLJHGPDLTINRQIGSSVFPEVQEDSKHA	120	
QY	121	VGORNGVETSVVLSDOEYARLQSIDPEGRKDFEFTGGRGAGHAMTVASDITEARQILI	180	
Db	121	VGORNGVETSVVLSDOEYARLQSIDPEGRKDFEFTGGRGAGHAMTVASDITEARQILI	180	
QY	181	ELLEPKGTGESGAGESKGVGELRBSNSGAEMTTETQTSSTSSLSRSDPKMLALGTAVT	240	
Db	181	ELLEPKGTGESGAGESKGVGELRBSNSGAEMTTETQTSSTSSLSRSDPKMLALGTAVT	240	
QY	241	GLIGLAATGIVQALALTPEDSPPTTDPAAASATEFARDDLTKEAFQNPNDOKNIDE	300	
Db	241	GLIGLAATGIVQALALTPEDSPPTTDPAAASATEFARDDLTKEAFQNPNDOKNIDE	300	
QY	301	LGNALPSGVLKDDVVANIEQAKAAGEAKQOAIENNAOAKKYDEQAKROEELKVSSG	360	
Db	301	LGNALPSGVLKDDVVANIEQAKAAGEAKQOAIENNAOAKKYDEQAKROEELKVSSG	360	
QY	361	AGYGSGLALITGGGIGVAVTALAHKRNQVVEQTTTTTTTTTSA7YENKPAANNPAQG	420	
Db	361	AGYGSGLALITGGGIGVAVTALAHKRNQVVEQTTTTTTTTTSA7YENKPAANNPAQG	420	
QY	421	NVDTPGSEDTMESRRSSMASTSTFFDTSISICTFVNQPADVKTSLHSDQVPTSNSNTSVQ	480	

DB	Sequence	Position
Db	421 NVDTPGSDTDTHESRSSMASTSTSTFDTSISICTVONPPADVTSI.LHDSQVPIPSNSTSVQ	480
Qy	481 NMGNITDSVVYSTIQHPPRDITDNGARLLGNPSAGIQSTYARLALSGGLRHMGGITGGSN	540
Db	481 NMGNITDSVVYSTIQHPPRDITDNGARLLGNPSAGIQSTYARLALSGGLRHMGGITGGSN	540
Qy	541 SAVNTSNNPAPGSHRFV 558	
Db	541 SAVNTSNNPAPGSHRFV 558	

```

RESULT 3
US-11-097-143-31017
; Sequence 31017, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; TITLE OF INVENTION: DROSOPHILA GENES.
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 31017
; LENGTH: 1300
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-11-097-143-31017

```

Query Match	6.6%	Score 187	DB 6	Length 1300
Best Local Similarity	20.4%	Pred. No. 0.003		
Matches	130	Conservative	74	Mismatches 245
				Indels 188
				Gaps 25
QY	5	NLGHNPVNNSIPAPPLPSQTDG--AGGRGQLINSTGPLGRALLFPVRNNSMADSDN	61	
DB	438	NSTGNSNSTMDSTGPFSETSTNGLVASGGAG--GATGAAWLPPP--SQOSTGK	488	
QY	62	RAS-----DVPLPVPMPRLA-----ASEITLNDGEVLHDGPDLLTRQ	102	
DB	489	EATAAVSLLEKKLPNVVSPFLTMELRKQKMTKYDAEMIMNAAVQOQH--	539	
QY	103	TGSSVFRVETQDEGKHLAVGQNGVETSVLSDQFARLQSIDPEGDKKVFYTGRCGAG	162	
DB	540	--OHFIHHHHHHHHHHGHASTGAEATAAVQOAMQKRG-----VGTGGAAG	588	
QY	163	HAMTVASDITEARQRILELLEPKQTGSKGAGEKGVGEKRESNGAENTTETQST	222	
DB	589	NAGATTSSVA-----AGAGEVNVGGRSTSLRKMRVNS-----TSSSI	627	
QY	223	SSLRSDEPKMLATGTVATGLIGLAATGIYQALALTPBEDSPFTTDPDAASATEIATRDQ	282	
DB	628	STASADEVY-----APVVAASISLPKAPVLMIRCKPAPDMALIALHQ	670	

QY 283 LTRKAFQNDQKONIDELGNALPSGVLDKDDVYANIEBQAKAGEAKQO-----A 333
DB 671 SOORQURRSEBQKEXLTD-----GSSD--TSEBQKX-----EQKQDQHLQPKMS 716
QY 334 IENNAQOKKYDEQOAKROBELKVSSGAGYGLSGALLIGGIGVAVTALHRRKQPEQOT 393
DB 717 LAEEPQPEKS-ERKQOEQOKRVYRNSAGRVGL-----VARLATMNNN-----1A 760
QY 394 TTTTNTTTSARTV---ENKPAANTPAQGVNDTPGSEDIMESRRSMASTSTFPDTS 450
DB 761 TTTNSSSSSKATTTTCNNHNSNSRINHSNLSSRUSVSKRKAPSEASSIPGSTS 820
QY 451 IGTVO-----NPYADVKTSLHDSQVPTSNSTSVQNMGNTDSVYSTIOH-----495
DB 821 EMQOQOATRRSCSPFTAYVKNLASFDPDPBPTQGIKEQKDBSVYTSVPVKQKRRRAA 880
QY 496 -----PRDPTDNGARL-IGNPSAGIQT-----YARLALSGG 527
DB 881 LAAASIHCEALGFPYTGSTGSRKRAQAGEPTTSCSTTISNVBPLKTPERLKLTLR 940
QY 528 LRH-----DMGGLTGGNSAVNTSNPPAPGSHR 556
DB 941 MKRSPLDEVIEBG--TSLNCGAGRG---APGSHR 971

RESULT 4
US-11-097-143-27426
; Sequence 27426, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27426
; LENGTH: 2586
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-11-097-143-27426

Query Match 6.4%; Score 182.5; DB 6; Length 2586;
Best Local Similarity 19.4%; Pred. No. 0.014;
Matches 113; Conservative 85; Mismatches 237; Indels 147; Gaps 19;

QY 113 QEDGKHIAVGQNGVETSVLSQEVARIQSIDP-----EGDKXFFVTGGRGAGHAMT 167
DB 183 QSSS-----STTTSSDEQOTTSSDPVVEVAQSS-----SNGDGNSTQSLTT 227
QY 168 VASDIPEARQRLLELPEKGTGESKAGESKGGELRESGAENTTEOTST-----222
DB 228 TTTTTS-----SDQGQSTSSDPVVEVSGTNGNSSTQSSSTTTTSSD 275
QY 223 ---SSLRDPKMLALGTATGILGLAATGIVQALALTEPDSPTTTPDPAASATETAT 279
DB 276 EGQTTSSDPVVEVAQSSSSNG-----DGNSTQSSSTTTT 310
QY 280 RQUTKEAFQNDQKONIDELGNALPSGVLDKDDVYANIEBQAKAGEAKQOAIENNAQ 339
DB 311 TTTTSSDGGQ-----STLLSDPVV-----EVSQGTNGNS 341
QY 340 AQKYDEQOAKROBELKVSSGA---GYGLSGLALIGGIGVAVTALHRRKQPEQOTTT 396
DB 342 TQSSSTTTTSSDEQOTTSSSDPVGEVAQSSSNGDGNST-----QSSTT 387
QY 397 TTTTTSARTVENKPAANTPAQGVNDTPGSEDIMESRRSMASTSTFPDTSIGTVON 456
DB 388 TTTTTSDDGQSTSSDPVVEVSGTNGN---STQSSSATTTTSSDEQOTTSSSD 444
QY 457 PYADV---KTSLHDSQVPTSNSTSVQNMGNTDSVYSTIOHPDPTD--NGARLGNP 511
DB 445 PVSEVAQSSSIGDGNSTQSTTTTSSDGGQSTSSDPVVEVSGTNGN---GNS 500
QY 512 SAGIOSYVARLALSGLRHDMGGLTGGNSAVNTSNPPAPG 553
DB 501 STQSSSTTTTSS-----DEQOTTSSDPVVEVAQSSSSNG 537

RESULT 5
US-09-946-374-243
; Sequence 243, Application US/09946374
; Publication No. US20030073129A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnovers, Luc
; APPLICANT: Baton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2830PICI
; CURRENT APPLICATION NUMBER: US/09/946,374
; CURRENT FILING DATE: 2001-09-04
; PRIOR APPLICATION NUMBER: 60/098716
; PRIOR FILING DATE: 1998-09-01
; PRIOR APPLICATION NUMBER: 60/098723
; PRIOR FILING DATE: 1998-09-01
; PRIOR APPLICATION NUMBER: 60/098749
; PRIOR FILING DATE: 1998-09-01
; PRIOR APPLICATION NUMBER: 60/098750
; PRIOR FILING DATE: 1998-09-01

PRIOR FILING DATE:	1998-09-23
PRIOR APPLICATION NUMBER:	60/101472
PRIOR FILING DATE:	1998-09-23
PRIOR APPLICATION NUMBER:	60/101474
PRIOR FILING DATE:	1998-09-23
PRIOR APPLICATION NUMBER:	60/101475
PRIOR FILING DATE:	1998-09-23
PRIOR APPLICATION NUMBER:	60/101476
PRIOR FILING DATE:	1998-09-23
PRIOR APPLICATION NUMBER:	60/101477
PRIOR FILING DATE:	1998-09-23
PRIOR APPLICATION NUMBER:	60/101479
PRIOR FILING DATE:	1998-09-23
PRIOR APPLICATION NUMBER:	60/101738
PRIOR FILING DATE:	1998-09-24
PRIOR APPLICATION NUMBER:	60/101741
PRIOR FILING DATE:	1998-09-24
PRIOR APPLICATION NUMBER:	60/101743
PRIOR FILING DATE:	1998-09-24
PRIOR APPLICATION NUMBER:	60/101915
PRIOR FILING DATE:	1998-09-24
PRIOR APPLICATION NUMBER:	60/101916
PRIOR FILING DATE:	1998-09-24
PRIOR APPLICATION NUMBER:	60/102207
PRIOR FILING DATE:	1998-09-29
PRIOR APPLICATION NUMBER:	60/102240
PRIOR FILING DATE:	1998-09-29
PRIOR APPLICATION NUMBER:	60/102307
PRIOR FILING DATE:	1998-09-29
PRIOR APPLICATION NUMBER:	60/102330
PRIOR FILING DATE:	1998-09-29
PRIOR APPLICATION NUMBER:	60/102331
PRIOR FILING DATE:	1998-09-29
PRIOR APPLICATION NUMBER:	60/102484
PRIOR FILING DATE:	1998-09-30
PRIOR APPLICATION NUMBER:	60/102487
PRIOR FILING DATE:	1998-09-30
PRIOR APPLICATION NUMBER:	60/102570
PRIOR FILING DATE:	1998-09-30
PRIOR APPLICATION NUMBER:	60/102571
PRIOR FILING DATE:	1998-09-30
PRIOR APPLICATION NUMBER:	60/102664
PRIOR FILING DATE:	1998-10-01
PRIOR APPLICATION NUMBER:	60/102687
PRIOR FILING DATE:	1998-10-01
PRIOR APPLICATION NUMBER:	60/102965
PRIOR FILING DATE:	1998-10-02
PRIOR APPLICATION NUMBER:	60/103258
PRIOR FILING DATE:	1998-10-06
PRIOR APPLICATION NUMBER:	60/103314
PRIOR FILING DATE:	1998-10-07
PRIOR APPLICATION NUMBER:	60/103315
PRIOR FILING DATE:	1998-10-07
PRIOR APPLICATION NUMBER:	60/103328
PRIOR FILING DATE:	1998-10-07
PRIOR APPLICATION NUMBER:	60/103449
PRIOR FILING DATE:	1998-10-06
PRIOR APPLICATION NUMBER:	60/103633
PRIOR FILING DATE:	1998-10-08
PRIOR APPLICATION NUMBER:	60/103678
PRIOR FILING DATE:	1998-10-08
PRIOR APPLICATION NUMBER:	60/103679
PRIOR FILING DATE:	1998-10-08
PRIOR APPLICATION NUMBER:	60/103711
PRIOR FILING DATE:	1998-10-08
PRIOR APPLICATION NUMBER:	60/104257
PRIOR FILING DATE:	1998-10-14

;; PRIOR APPLICATION NUMBER: 60/104987
;; PRIOR FILING DATE: 1998-10-20
;; PRIOR APPLICATION NUMBER: 60/105000
;; PRIOR FILING DATE: 1998-10-20
;; PRIOR APPLICATION NUMBER: 60/105002
;; PRIOR FILING DATE: 1998-10-20
;; PRIOR APPLICATION NUMBER: 60/105104
;; PRIOR FILING DATE: 1998-10-21
;; PRIOR APPLICATION NUMBER: 60/105169
;; PRIOR FILING DATE: 1998-10-22
;; PRIOR APPLICATION NUMBER: 60/105266
;; PRIOR FILING DATE: 1998-10-22
;; PRIOR APPLICATION NUMBER: 60/105693
;; PRIOR FILING DATE: 1998-10-26
;; PRIOR APPLICATION NUMBER: 60/105694
;; PRIOR FILING DATE: 1998-10-26
;; PRIOR APPLICATION NUMBER: 60/105807
Query Match 6.3%; Score 180; DB 3; Length 596;
Best Local Similarity 19.2%; Pred. No. 0.0031;
Matches 109; Conservative 90; Mismatches 272; Indels 98; Gaps 19;
QY 4 GNLGNPNVNSIPAPLPISQTDGAGRGQLINSTGPGSRALFTPVNSMADSGDNRA 63
DB 13 GLILHLEAATNS-----NETSTANTSSVSISSG-----ASTATNSG 49
QY 64 SDVPGLPVPMRLAASEITLNDGFEVLHDHGPLDTLNRQIGSSV-FRVEIOEGKHIAVG 122
DB 50 SSTSSGVSSTATISGSSVTSN-GVSIY-TNSEPHITSSGISTATNSEFTASSGISIATN 107
QY 123 ORNGVETS--VVLSDOEYARLOSIDPEGKOKFVETGGRGAGHAMVTASDITEARQIL 180
DB 108 SESSTSSGASTATNSE---SSTPSSGASTVTSNSSSVTSAGASTATNSESTVSSRPS 163
QY 181 ELLEPKGTESKAGESKGVGELRESNGALETTEOTSTSTSLRSDPMLALGTVAT 240
DB 164 TATNSSSTLSSGASTATN-SDSSTSSGASTATNSESTSSG-----ASTAT 211
QY 241 GLIGLATGIVQALITPEPDSPTTDPDAASATETATRDQITKEAFONPDOKVNIIDE 300
DB 212 N-----SESTVSSRASTATNSESTT---SSGASTATNSESTRTSSGAGATATNSESTTS 264
QY 301 LG-----NAIPSGVLKDVVANIIEGAKAAGEBAKQOALENNAAQKXDEQOAKR 351
DB 265 SGASTATNSDSSVSSGA---STATNSESTSSGAST-----ATN 302
QY 352 OEBLKYSAGAGYGL-SCALILGGIGVAVTAAHLRKNQPEQTTTTTTTTTTSARTVEN 410
DB 303 SESSTSSGASTATNSDSSSTSSGAGATATNSESTVSSGISITVTSSESTPSSGANTATN 362
QY 411 KPAAANTPAQGNVDTPESEDTMESRRSRMASTSTFFDTSSIGTVONPVADKTS----- 464
DB 363 SESSTSSGANTAT-NSSESTVSSGASTATNSESTSSGVSSTATNSESTSSGASTAT 421
QY 465 LHDQVPTNSNMTSVONMGNTOSVVVSTIOHPRDITDNCARLLGPSAGIOSTYARLAL 544
DB 422 NSDSSSTSSSEASTATNSESTVSSGISITVTSSESTSSGANTATATNSGSVTS----- 474
QY 525 SGGLRHDMGJLTCGNSNAVTNSNPPAPG 553
DB 475 ACSGTAALTCGHTTSHS-STAIVSEAKPG 502
RESULT 6
US-10-006-867-100
; Sequence 100, Application US/10006867
; Publication No. US20020119130A1
; GENERAL INFORMATION:
; APPLICANT: Eaton, Dan L.
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gerlitsen, Mary E.
; APPLICANT: Goddard, Audrey J.
; APPLICANT: Godowski, Paul J.

;; APPLICANT: Grimaldi, Christopher J.
;; APPLICANT: Gurney, Austin L.
;; APPLICANT: Watanabe, Colin K.
;; APPLICANT: Wood, William I.
;; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;; FILE REFERENCE: P323ORIC1
;; CURRENT APPLICATION NUMBER: US/10/006, 867
;; CURRENT FILING DATE: 2001-12-06
;; PRIOR APPLICATION NUMBER: 60/063435
;; PRIOR FILING DATE: 1997-10-29
;; PRIOR APPLICATION NUMBER: 60/064215
;; PRIOR FILING DATE: 1997-10-29
;; PRIOR APPLICATION NUMBER: 60/082797
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/083495
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/085579
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/087759
;; PRIOR FILING DATE: 1998-06-02
;; PRIOR APPLICATION NUMBER: 60/088021
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088029
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088030
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088734
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088740
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088811
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088824
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088825
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088863
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/089105
;; PRIOR FILING DATE: 1998-06-12
;; PRIOR APPLICATION NUMBER: 60/089514
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089653
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089952
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/090246
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090444
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090688
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090696
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090862
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/091628
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/096012
;; PRIOR FILING DATE: 1998-08-10
;; PRIOR APPLICATION NUMBER: 60/096757
;; PRIOR FILING DATE: 1998-08-17
;; PRIOR APPLICATION NUMBER: 60/096949
;; PRIOR FILING DATE: 1998-08-18
;; PRIOR APPLICATION NUMBER: 60/096959
;; PRIOR FILING DATE: 1998-08-18
;; PRIOR APPLICATION NUMBER: 60/097954
;; PRIOR FILING DATE: 1998-08-26
;; PRIOR APPLICATION NUMBER: 60/097971
;; PRIOR FILING DATE: 1998-08-26
;; PRIOR APPLICATION NUMBER: 60/097979
;; PRIOR FILING DATE: 1998-08-26

PRIOR APPLICATION NUMBER: 60/098749
PRIOR FILING DATE: 1998-09-01
PRIOR APPLICATION NUMBER: 60/099741
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099763
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099792
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099812
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099815
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/100627
PRIOR FILING DATE: 1998-09-16
PRIOR APPLICATION NUMBER: 60/100662
PRIOR FILING DATE: 1998-09-16
PRIOR APPLICATION NUMBER: 60/100683
PRIOR FILING DATE: 1998-09-17
PRIOR APPLICATION NUMBER: 60/100684
PRIOR FILING DATE: 1998-09-17
PRIOR APPLICATION NUMBER: 60/100930
PRIOR FILING DATE: 1998-09-17
PRIOR APPLICATION NUMBER: 60/101279
PRIOR FILING DATE: 1998-09-22
PRIOR APPLICATION NUMBER: 60/101475
PRIOR FILING DATE: 1998-09-23
PRIOR APPLICATION NUMBER: 60/101738
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/101743
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/101916
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/102570
PRIOR FILING DATE: 1998-09-30
PRIOR APPLICATION NUMBER: 60/103449
PRIOR FILING DATE: 1998-10-06
PRIOR APPLICATION NUMBER: 60/103678
PRIOR FILING DATE: 1998-10-08
PRIOR APPLICATION NUMBER: 60/103679
PRIOR FILING DATE: 1998-10-08
PRIOR APPLICATION NUMBER: 60/103711
PRIOR FILING DATE: 1998-10-08
PRIOR APPLICATION NUMBER: 60/105000
PRIOR FILING DATE: 1998-10-20
PRIOR APPLICATION NUMBER: 60/105002
PRIOR FILING DATE: 1998-10-20
PRIOR APPLICATION NUMBER: 60/105881
PRIOR FILING DATE: 1998-10-27
PRIOR APPLICATION NUMBER: 60/106030
PRIOR FILING DATE: 1998-10-28
PRIOR APPLICATION NUMBER: 60/106464
PRIOR FILING DATE: 1998-10-30
PRIOR APPLICATION NUMBER: 60/106856
PRIOR FILING DATE: 1998-11-03
PRIOR APPLICATION NUMBER: 60/108807
PRIOR FILING DATE: 1998-11-17
PRIOR APPLICATION NUMBER: 60/112419
PRIOR FILING DATE: 1998-12-15
PRIOR APPLICATION NUMBER: 60/112422
PRIOR FILING DATE: 1998-12-15
PRIOR APPLICATION NUMBER: 60/112853
PRIOR FILING DATE: 1998-12-16
PRIOR APPLICATION NUMBER: 60/113011
PRIOR FILING DATE: 1998-12-16
PRIOR APPLICATION NUMBER: 60/112854
PRIOR FILING DATE: 1998-12-16
PRIOR APPLICATION NUMBER: 60/113300
PRIOR FILING DATE: 1998-12-22
PRIOR APPLICATION NUMBER: 60/113408
PRIOR FILING DATE: 1998-12-22
PRIOR APPLICATION NUMBER: 60/113430
PRIOR FILING DATE: 1998-12-23
PRIOR APPLICATION NUMBER: 60/113621

PRIOR FILING DATE: 1998-12-23
PRIOR APPLICATION NUMBER: 60/114223
PRIOR FILING DATE: 1998-12-30
PRIOR APPLICATION NUMBER: 60/115614
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/116527
PRIOR FILING DATE: 1999-01-20
PRIOR APPLICATION NUMBER: 60/116843
PRIOR FILING DATE: 1999-01-22
PRIOR APPLICATION NUMBER: 60/119285
PRIOR FILING DATE: 1999-02-09
PRIOR APPLICATION NUMBER: 60/119287
PRIOR FILING DATE: 1999-02-09
PRIOR APPLICATION NUMBER: 60/119525
PRIOR FILING DATE: 1999-02-10
PRIOR APPLICATION NUMBER: 60/119549
PRIOR FILING DATE: 1999-02-10
PRIOR APPLICATION NUMBER: 60/120014
PRIOR FILING DATE: 1999-02-11
PRIOR APPLICATION NUMBER: 60/129122
PRIOR FILING DATE: 1999-04-13
PRIOR APPLICATION NUMBER: 60/129674
PRIOR FILING DATE: 1999-04-16
PRIOR APPLICATION NUMBER: 60/131291
PRIOR FILING DATE: 1999-04-27
PRIOR APPLICATION NUMBER: 60/138387
PRIOR FILING DATE: 1999-06-09
PRIOR APPLICATION NUMBER: 60/144791
PRIOR FILING DATE: 1999-07-20
PRIOR APPLICATION NUMBER: 60/169495
PRIOR FILING DATE: 1999-12-07
PRIOR APPLICATION NUMBER: 60/175481
PRIOR FILING DATE: 2000-01-11
PRIOR APPLICATION NUMBER: 60/191007
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/199397
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/380139
PRIOR FILING DATE: 1998-08-25
PRIOR APPLICATION NUMBER: 09/311832
PRIOR FILING DATE: 1999-05-14
PRIOR APPLICATION NUMBER: 09/380137
PRIOR FILING DATE: 1999-08-25
PRIOR APPLICATION NUMBER: 09/380138
PRIOR FILING DATE: 1999-08-25
PRIOR APPLICATION NUMBER: 09/380142

Query Match 6.3%; Score 180; DB 4; Length 596;
Best Local Similarity 19.2%; Pred. No. 0.0031;
Matches 109; Conservative 90; Mismatches 272; Indels 98; Gaps 19;

QY 4 GNLGNPNVNNISIPAPPLPSQTDGAGRGQLINSTGLGRALFTPRNMAPOGDURA 63
DB 13 GLLHLLEAATNS-----NETSTANTSSVSSG-----ASTATNSG 49
QY 64 SDVPLPVNPKRLASBITLNDGFEVLHDHGPLDTLNRQIGSV-FVETQEDGKHLAVG 122
DB 50 SSVTSSGVSTATISSSTSN-GVSI-VTNEFHRTSGISTATNSESTSSGSIATN 107
QY 123 ORNGVETS--VVLSDQEVARLOSIDPEGKDFVFTGGRGAGHAMVTVASDITEARQIL 180
DB 108 SSSSTSSGASTATNSE-----STPSSGASTVTVNSGSVTVSSGASTATNSESTVSRAS 163
QY 181 ELLEPKTGESKKGAGESGVGELEKRSNGAENTTETOTSTSTSSLRSPKMLALGYAT 240
DB 164 TATNSESSTLSSGASTATN-SDSSTSSGASTATNSESTTSSG-----ASTAT 211
QY 241 GLIGLAATGIYOALATPEPDPPTTDPDAASATETATRPOLTYEATFQNPONKYNIDE 300
DB 212 N-----SESTVSSRSTATNSESTT---SSGASTATNSEBRTTNGAGTATNSESTTS 264
QY 301 LG-----NALPSGYLKDDVVANIEQAKAAGEAKQQAIENTNAQAQKKYDEQAKR 351

```
Db 265 SGASTATNSDSTVSSGA---STATNSESSTSSGAST-----ATN 302
QY 352 QEBLKXSSGAGYGL-SGALLGGIGVAVTAALHRKQVPEQTTTTTTTTTSARVEN 410
Db 303 SESSTTSSGASTATNSDSTSSGAGATATNSESSTVSSGISTVNSSSPSSGANTATN 362
QY 411 KPANNTPAQGNVDPDTPGSEDTMESRRSSMASTSTFFDTSSIGTVQNPYAVKTS-----464
Db 363 SESSTTSSGANTAT-NSSESSTVSSGASTATNSESSTSSGASTATNSESSTSSGASTAT 421
QY 465 LHDQVPTNSNTSVONMGNTDSVYSTIOHPPRDTTNGARLLGNPSAGIOSYARLAL 524
Db 422 NSDSTSSSEASATATNSESSTVSSGISTVNSESSTSSGANTATNSESSTVSS-474
QY 525 SGLRDMGGLTGSNSAVTNSNPPAPG 553
Db 475 AGSGTAALTGMTTSHS-STAVSEAKPG 502

RESULT 7
US-10-052-586-310
; Sequence 310, Application US/10052586
; Publication No. US20020127584A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C1
; CURRENT APPLICATION NUMBER: US/10/052, 586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063564
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063734
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/063870
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066120
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/066466
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/066772

; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/069335
; PRIOR FILING DATE: 1997-12-11
; PRIOR APPLICATION NUMBER: 60/069425
; PRIOR FILING DATE: 1997-12-12
; PRIOR APPLICATION NUMBER: 60/069870
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/068017
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/077450
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 60/077632
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077649
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/078886
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078939
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079664
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079786
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/080107
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080194
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080327
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080333
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/081049
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081070
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081195
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081838
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/082568
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082569
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082704
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082797
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/083495
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083496
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083499
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083559
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/084366
; PRIOR FILING DATE: 1998-05-05
; PRIOR APPLICATION NUMBER: 60/084414
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084639
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084640
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084643
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
```

PRIOR APPLICATION NUMBER: 60/085582
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085700
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/086023
PRIOR FILING DATE: 1998-05-18
PRIOR APPLICATION NUMBER: 60/086392
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/086486
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/087098
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087208
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088722
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088740
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088811
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088825
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088863
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089090
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653

PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089908
Query Match 6.3%; Score 180; DB 4; Length 596;
Best Local Similarity 19.2%; Pred. No. 0.0031;
Matches 109; Conservative 90; Mismatches 272; Indels 98; Gaps 19;
QY 4 GNLGNPNVNNISIPAPPLPSQTDGAGRGQLINSTGLSRALPTPVNNSMADSGDRA 63
DB 13 GLLHLEAATNS-----NETSANTGSSVSSG-----ASTATNSG 49
QY 64 SDVPLPVNPMRLAASEITLNDGEVLHDHPLDTLNRQISSV-FRVEYEDGHIIVG 122
DB 50 SSVTSGVSTATISSSVTSN-GVSIIV-TNSEFTTSSGISTATSEFSTASGSIATN 107
QY 123 QANGVETS--VLSQGEFARLQSIDPEKDKFVFCGAGGAGHAMVATVASDITERRQIL 180
DB 108 SSSSTSSGASTATNSE---SSTPSSGASTVTVNSSSVTSSTGASTATNSESSTSSRAS 163
QY 181 ELLEPKGTGESKAGESKVGELRESNGAENTTETOTSTSTSLRSPKMLALGTVA 240
DB 164 TATNSESSTLSSGASTATN-SDSSTSSGASTATNSESSTSSG-----ASTAT 211
QY 241 GLIGLAATGIYQALALTPEDPSPTTDDAASATETARDQLTKAEPNDQKNIDE 300
DB 212 N---SSSTVSSRASTATNSESST---SSGASTATNSESRTTNGAGTATNSESSTTS 264
QY 301 LG-----NAIPSSGILKDDVANNIEQAKAGEAKQAIAENNAQAOKKYDEQOAKR 351
DB 265 SGASTATNSDSSTVSSGA---STATNSESSTSSGAST-----ATN 302
QY 352 QEELKVSAGAGYGL-SCALLIGGIGVAVTAALHRKNQVPEQTTTTTTTTTSARVFN 410
DB 303 SSSSTSSGASTATNSDSSTSSGAGTATNSESSTSSGISTVTVNSESSTSSGANTATN 362
QY 411 KPAANTPAQGVNTPGSSDPTMESRRSSMASTSTFFDTSSIGTVQNPYADVKT----- 464
DB 363 SSSSTSSGANTAT-NSSESTVSSGASTATNSESSTSSGASTATNSESSTSSGASTAT 421
QY 465 LHDQVPTNSNTSVQNMGNNDVSVYTIQHPRPRTTNGARLGNPAGIOSTARLAL 524
DB 422 NSDSSTSSSEASTATNSESSTVSSGISTVTVNSESSTSSGANTATNSESSTSVTS----- 474
QY 525 SGGLRHDMGGLTGSGNSAVNTSNPPAG 553
DB 475 AGSGTALTGHTTTHSHA-STAVSEAKPG 502
RESULT 8
US-10-063-547-100
Sequence 100, Application US/10063547
Publication No. US20020182638A1
GENERAL INFORMATION:
APPLICANT: Eaton, Dan L.
APPLICANT: Filvaroff, Ellen
APPLICANT: Geritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Gurney, Austin L.
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3230R1C1
CURRENT APPLICATION NUMBER: US/10/063,547
CURRENT FILING DATE: 2002-05-02
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 170
SEQ ID NO 100
LENGTH: 596
TYPE: PRT
ORGANISM: Homo Sapien

US-10-174-590-310

Query Match 6.3%; Score 180; DB 4; Length 596;

Best Local Similarity 19.2%; Pred. No. 0.0031;

Matches 109; Conservative 90; Mismatches 272; Indels 98; Gaps 19;

```
4 GNLGNPNVNNISIPAPPLPSQTDGAGRGQLINSTGPIGSRALFTPVNNSMADSGDRA 63
13 GLHLLEAATNS-----NETSANTGSSVSSG-----ASTATNSG 49
64 SDVPLPVNPKRLAASEITLNDGFEVLHDHGPLDILNRQIGSV-FRVETOEDGKHIAVG 122
50 SSVTSSGVSATFISSSVTSN-GVSIY-TNSEPHITSSGISTATNSESTASSGSIATN 107
123 QNRGVETS-VVLSQDEYARLOSIDPEKDKFVFGGKGAGHAMVYASDITTEARQIL 180
108 SSSSTSSGASTATNSE-----STPPSGASTVTVNNGSSVTSNGASTATNSSTVSSRAS 163
181 ELLEPKGTGSGKAGESKGVGELRNSNGAENTETOTSTSTSLRSDPKMLALGTVAT 240
164 TATNSESTLSSGASTATN-SDSSTSSGASTATNSESTSSG-----ASTAT 211
241 GLIGLAATGIYQALALTPEDPSPTTTPDPAASATEFATRDQLTKEAFQNPDKVNIDE 300
212 N-----SESSIVSSRASTATNSESTT---SSGASTATNSESTRTSNGAGTATNSSESTTS 264
301 LG-----NAIPSGVLKDDVYANIEEQAKAGEEAKQQAIEENNAQAKKYDEQAKR 351
265 SGASTATNSDSSTVSSGA---STATNSSESTSSGAST-----ATN 302
352 QEBLKVSSGAGYGL-SCALLIGGIGIVAVTALHKKNPVEQTTTTTTTTTSARTVEN 410
303 SSSSTSSGASTATNSDSSTSSGAGTATNSSESTVSSGISTVTVNSESTSSGANTATN 362
411 KPANTPAQGVNDTPGSEDTMESRRSSMASTSTFFDTSSIGTVONPYADVKT----- 464
363 SSSSTSSGANTAT-NSSESTVSSGASTATNSSESTSSGASTATNSSESTSSGASTAT 421
465 LHDQVPTNSNTSVQNMGNITDSVYSTIQHPRTTNGARLLGNPAGIOSYARLAL 524
422 NSDSSTSSSEASTATNSSESTVSSGISTVTVNSESTSSGANTATNSGSSVTS----- 474
525 SGGLRHDMGGLTGSGNSAVNTSNPPAPG 553
475 AGSGTALTGMHTTSHS-STAVSEAKPG 502
```

RESULT 11
US-10-176-758-310

; Sequence 310, Application US/10176758

; Publication No. US20030008353A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Chen, Jian

; APPLICANT: Desnoyers, Luc

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Pan, James

; APPLICANT: Smith, Victoria

; APPLICANT: Watanabe, Colin K.

; APPLICANT: Wood, William I.

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; FILE REFERENCE: P3430R1C104

; CURRENT APPLICATION NUMBER: US/10/176,758

; CURRENT FILING DATE: 2002-06-21

; Prior Application removed - See File Wrapper or Palm

; NUMBER OF SEQ ID NOS: 612

; SEQ ID NO 310

; LENGTH: 596

; TYPE: PRT

; ORGANISM: Homo Sapien
US-10-176-758-310

Query Match 6.3%; Score 180; DB 4; Length 596;

Best Local Similarity 19.2%; Pred. No. 0.0031;

Matches 109; Conservative 90; Mismatches 272; Indels 98; Gaps 19;

```
4 GNLGNPNVNNISIPAPPLPSQTDGAGRGQLINSTGPIGSRALFTPVNNSMADSGDRA 63
13 GLHLLEAATNS-----NETSANTGSSVSSG-----ASTATNSG 49
64 SDVPLPVNPKRLAASEITLNDGFEVLHDHGPLDILNRQIGSV-FRVETOEDGKHIAVG 122
50 SSVTSSGVSATFISSSVTSN-GVSIY-TNSEPHITSSGISTATNSESTASSGSIATN 107
123 QNRGVETS-VVLSQDEYARLOSIDPEKDKFVFGGKGAGHAMVYASDITTEARQIL 180
108 SSSSTSSGASTATNSE-----STPPSGASTVTVNNGSSVTSNGASTATNSSTVSSRAS 163
181 ELLEPKGTGSGKAGESKGVGELRNSNGAENTETOTSTSTSLRSDPKMLALGTVAT 240
164 TATNSESTLSSGASTATN-SDSSTSSGASTATNSESTSSG-----ASTAT 211
241 GLIGLAATGIYQALALTPEDPSPTTTPDPAASATEFATRDQLTKEAFQNPDKVNIDE 300
212 N-----SESSIVSSRASTATNSESTT---SSGASTATNSESTRTSNGAGTATNSSESTTS 264
301 LG-----NAIPSGVLKDDVYANIEEQAKAGEEAKQQAIEENNAQAKKYDEQAKR 351
265 SGASTATNSDSSTVSSGA---STATNSSESTSSGAST-----ATN 302
352 QEBLKVSSGAGYGL-SCALLIGGIGIVAVTALHKKNPVEQTTTTTTTTTSARTVEN 410
303 SSSSTSSGASTATNSDSSTSSGAGTATNSSESTVSSGISTVTVNSESTSSGANTATN 362
411 KPANTPAQGVNDTPGSEDTMESRRSSMASTSTFFDTSSIGTVONPYADVKT----- 464
363 SSSSTSSGANTAT-NSSESTVSSGASTATNSSESTSSGASTATNSSESTSSGASTAT 421
465 LHDQVPTNSNTSVQNMGNITDSVYSTIQHPRTTNGARLLGNPAGIOSYARLAL 524
422 NSDSSTSSSEASTATNSSESTVSSGISTVTVNSESTSSGANTATNSGSSVTS----- 474
525 SGGLRHDMGGLTGSGNSAVNTSNPPAPG 553
475 AGSGTALTGMHTTSHS-STAVSEAKPG 502
```

RESULT 12
US-10-175-737-310

; Sequence 310, Application US/10175737

; Publication No. US20030013153A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Chen, Jian

; APPLICANT: Desnoyers, Luc

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Pan, James

; APPLICANT: Smith, Victoria

; APPLICANT: Watanabe, Colin K.

; APPLICANT: Wood, William I.

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; FILE REFERENCE: P3430R1C50

; CURRENT APPLICATION NUMBER: US/10/175,737

; CURRENT FILING DATE: 2002-06-19

; Prior Application removed - See File Wrapper or Palm

; NUMBER OF SEQ ID NOS: 612

; SEQ ID NO 310

; LENGTH: 596

```

; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-175-737-310

Query Match
Best Local Similarity 19.2%; Score 180; DB 4; length 596;
Matches 109; Conservative 90; Mismatches 272; Indels 98; Gaps 19;

QY 4 GNLGHNPNVNNNSIPPAPELPBSCQTDGAGRGQLNSTGPGLSRALFTPVNRSMADSGDNRA 63
DB 13 GLLHLLEAATNS-----NETSTANTGSSVSISSG-----ASTATNSG 49
QY 64 SDVPGLPVPMRLAASEITLNDGFEVLHDGPLDTLNROIGSSV-FRVETQEDGKIAYG 122
DB 50 SSTSSGCVSTATISGSSVTSN-GVSTV-TNSEPHITSSGISTATNSEPSTASSGISIATN 107
QY 123 QNRNGVETS--VLSDQEVARLOSIDPEGKDFVFTGGRGAGHAMVTVASDITEARQRL 180
DB 108 SESSSTSSGASTATNBE-----STPSSGASTVTNCGSVTSSGASTATNSESSTVSRRAS 163
QY 181 ELLEPKGTGEBKAGSKGVGELRESNGAENTETQTSTSTSLSDPKMLALGTVAAT 240
DB 164 TATNSESSTLSSGASTATN-SDSSTTSSGASTATNSESSTTSSG-----ASTAT 211
QY 241 GLIGLAATGIVQALALTPPEPDSPTTDPDAAASATETATRDOLTKAFAQNPQNOXNIDE 300
DB 212 N-----SESTVSSRASTATNSESSTT---SSGASTATNSESRTTSSGAGTATNSESSTTS 264
QY 301 LG-----NAIPSGVLKDVVANIIEQAKAGAEAKQAQIENNAQAKKYDEQOAKR 351
DB 265 SGASTATNDSSTVSSGA-----STATNSESSTTSSGAST-----ATN 302
QY 352 QRELKXSSGAGYGL-SCALLIGGIGVAYTAAALHRNQVEQTTTTTTTTTTSANTVEN 410
DB 303 SESSSTSSGASTATNDSSTSSGAGTATNSESSTVSSGISTVTNSESSTSSGANTATN 362
QY 411 KPANNTPAQANVTPESEPTMESRRSSMASTSTFPDTSIGTVQNPVADVKTSS----- 464
DB 363 SSSSTTSSGANTAT-NSESTVSSGASTATNSESSTTSSGCVSTATNSESSTTSSGASTAT 421
QY 465 LHDQVPTSNSTSVQNMGNNTDSVVYSTIQHPRDITDNGALGNPAGIOSTVARLAL 524
DB 422 NDSSTTSSSEASTATNSESSTVSSGISTVTNSESSTTSSGANTATNSSGSVTS----- 474
QY 525 SGGLRHDMGGLTGGSNSAVNTSNNPPAPG 553
DB 475 AGSGTAALTGMMHTTSHA-STAVSEAKPG 502

RESULT 13
US-10-063-616-100
; Sequence 100, Application US/10063616
; Publication No. US20030013855A1
; GENERAL INFORMATION:
; APPLICANT: Baton, Dan L.
; APPLICANT: Filvarcoff, Ellen
; APPLICANT: Gerltisen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3230R1C1
; CURRENT APPLICATION NUMBER: US/10/063, 616
; PRIOR APPLICATION REMOVED - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 100
; LENGTH: 596
; TYPE: PRT
```

```

; ORGANISM: Homo Sapien
US-10-063-616-100

Query Match
Best Local Similarity 19.2%; Score 180; DB 4; length 596;
Matches 109; Conservative 90; Mismatches 272; Indels 98; Gaps 19;

QY 4 GNLGHNPNVNNNSIPPAPELPBSCQTDGAGRGQLNSTGPGLSRALFTPVNRSMADSGDNRA 63
DB 13 GLLHLLEAATNS-----NETSTANTGSSVSISSG-----ASTATNSG 49
QY 64 SDVPGLPVPMRLAASEITLNDGFEVLHDGPLDTLNROIGSSV-FRVETQEDGKIAYG 122
DB 50 SSTSSGCVSTATISGSSVTSN-GVSTV-TNSEPHITSSGISTATNSEPSTASSGISIATN 107
QY 123 QNRNGVETS--VLSDQEVARLOSIDPEGKDFVFTGGRGAGHAMVTVASDITEARQRL 180
DB 108 SESSSTSSGASTATNBE-----STPSSGASTVTNCGSVTSSGASTATNSESSTVSRRAS 163
QY 181 ELLEPKGTGEBKAGSKGVGELRESNGAENTETQTSTSTSLSDPKMLALGTVAAT 240
DB 164 TATNSESSTLSSGASTATN-SDSSTTSSGASTATNSESSTTSSG-----ASTAT 211
QY 241 GLIGLAATGIVQALALTPPEPDSPTTDPDAAASATETATRDOLTKAFAQNPQNOXNIDE 300
DB 212 N-----SESTVSSRASTATNSESSTT---SSGASTATNSESRTTSSGAGTATNSESSTTS 264
QY 301 LG-----NAIPSGVLKDVVANIIEQAKAGAEAKQAQIENNAQAKKYDEQOAKR 351
DB 265 SGASTATNDSSTVSSGA-----STATNSESSTTSSGAST-----ATN 302
QY 352 QRELKXSSGAGYGL-SCALLIGGIGVAYTAAALHRNQVEQTTTTTTTTTTSANTVEN 410
DB 303 SESSSTSSGASTATNDSSTSSGAGTATNSESSTVSSGISTVTNSESSTSSGANTATN 362
QY 411 KPANNTPAQANVTPESEPTMESRRSSMASTSTFPDTSIGTVQNPVADVKTSS----- 464
DB 363 SSSSTTSSGANTAT-NSESTVSSGASTATNSESSTTSSGCVSTATNSESSTTSSGASTAT 421
QY 465 LHDQVPTSNSTSVQNMGNNTDSVVYSTIQHPRDITDNGALGNPAGIOSTVARLAL 524
DB 422 NDSSTTSSSEASTATNSESSTVSSGISTVTNSESSTTSSGANTATNSSGSVTS----- 474
QY 525 SGGLRHDMGGLTGGSNSAVNTSNNPPAPG 553
DB 475 AGSGTAALTGMMHTTSHA-STAVSEAKPG 502

RESULT 14
US-10-174-581-310
; Sequence 310, Application US/10174581
; Publication No. US20030017540A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C41
; CURRENT APPLICATION NUMBER: US/10/174, 581
; PRIOR APPLICATION NUMBER: 10/052586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
```

PRIOR APPLICATION NUMBER: 60/059266	PRIOR APPLICATION NUMBER: 60/062250	PRIOR APPLICATION NUMBER: 60/063120	PRIOR APPLICATION NUMBER: 60/063421	PRIOR APPLICATION NUMBER: 60/063486	PRIOR APPLICATION NUMBER: 60/063540	PRIOR APPLICATION NUMBER: 60/063541	PRIOR APPLICATION NUMBER: 60/063544	PRIOR APPLICATION NUMBER: 60/063564	PRIOR APPLICATION NUMBER: 60/063734	PRIOR APPLICATION NUMBER: 60/063870	PRIOR APPLICATION NUMBER: 60/064103	PRIOR APPLICATION NUMBER: 60/065311	PRIOR APPLICATION NUMBER: 60/065311	PRIOR APPLICATION NUMBER: 60/066120	PRIOR APPLICATION NUMBER: 60/066466	PRIOR APPLICATION NUMBER: 60/066772	PRIOR APPLICATION NUMBER: 60/069335	PRIOR APPLICATION NUMBER: 60/069425	PRIOR APPLICATION NUMBER: 60/069870	PRIOR APPLICATION NUMBER: 60/068017	PRIOR APPLICATION NUMBER: 1997-12-18	PRIOR APPLICATION NUMBER: 60/07450	PRIOR APPLICATION NUMBER: 60/07632	PRIOR APPLICATION NUMBER: 60/077649	PRIOR APPLICATION NUMBER: 1998-03-11	PRIOR APPLICATION NUMBER: 60/078886	PRIOR APPLICATION NUMBER: 60/078939	PRIOR APPLICATION NUMBER: 1998-03-20	PRIOR APPLICATION NUMBER: 60/079654	PRIOR APPLICATION NUMBER: 1998-03-27	PRIOR APPLICATION NUMBER: 60/079786	PRIOR APPLICATION NUMBER: 60/080107	PRIOR APPLICATION NUMBER: 1998-03-31	PRIOR APPLICATION NUMBER: 60/080194	PRIOR APPLICATION NUMBER: 1998-03-31	PRIOR APPLICATION NUMBER: 60/080327	PRIOR APPLICATION NUMBER: 1998-04-01	PRIOR APPLICATION NUMBER: 60/080333	PRIOR APPLICATION NUMBER: 60/081049	PRIOR APPLICATION NUMBER: 1998-04-08	PRIOR APPLICATION NUMBER: 60/081070	PRIOR APPLICATION NUMBER: 60/081195	PRIOR APPLICATION NUMBER: 1998-04-09	PRIOR APPLICATION NUMBER: 60/081838	PRIOR APPLICATION NUMBER: 1998-04-15	PRIOR APPLICATION NUMBER: 60/082568	PRIOR FILING DATE: 1998-04-21	PRIOR APPLICATION NUMBER: 60/082569
PRIOR FILING DATE: 1997-03-18	PRIOR FILING DATE: 1997-10-17	PRIOR FILING DATE: 1997-10-24	PRIOR FILING DATE: 1997-10-24	PRIOR FILING DATE: 1997-10-24	PRIOR FILING DATE: 1997-10-28	PRIOR FILING DATE: 1997-10-28	PRIOR FILING DATE: 1997-10-28	PRIOR FILING DATE: 1997-10-28	PRIOR FILING DATE: 1997-10-29	PRIOR FILING DATE: 1997-11-24	PRIOR FILING DATE: 1997-11-21	PRIOR FILING DATE: 1997-11-13	PRIOR FILING DATE: 1997-11-24	PRIOR FILING DATE: 1997-11-24	PRIOR FILING DATE: 1997-12-11	PRIOR FILING DATE: 1997-12-11	PRIOR FILING DATE: 1997-12-11	PRIOR FILING DATE: 1997-12-11	PRIOR FILING DATE: 1997-12-17	PRIOR FILING DATE: 1997-12-17	PRIOR FILING DATE: 1997-12-18	PRIOR FILING DATE: 1998-03-10	PRIOR FILING DATE: 1998-03-10	PRIOR FILING DATE: 1998-03-11	PRIOR FILING DATE: 1998-03-11	PRIOR FILING DATE: 1998-03-11	PRIOR FILING DATE: 1998-03-20	PRIOR FILING DATE: 1998-03-20	PRIOR FILING DATE: 1998-03-27	PRIOR FILING DATE: 1998-03-27	PRIOR FILING DATE: 1998-03-27	PRIOR FILING DATE: 1998-03-27	PRIOR FILING DATE: 1998-03-31	PRIOR FILING DATE: 1998-03-31	PRIOR FILING DATE: 1998-03-31	PRIOR FILING DATE: 1998-03-31	PRIOR FILING DATE: 1998-03-31	PRIOR FILING DATE: 1998-03-31	PRIOR FILING DATE: 1998-04-01	PRIOR FILING DATE: 1998-04-01	PRIOR FILING DATE: 1998-04-01	PRIOR FILING DATE: 1998-04-08	PRIOR FILING DATE: 1998-04-08	PRIOR FILING DATE: 1998-04-08	PRIOR FILING DATE: 1998-04-09	PRIOR FILING DATE: 1998-04-15	PRIOR FILING DATE: 1998-04-21	PRIOR APPLICATION NUMBER: 60/082569

1	PRIOR FILING DATE: 1998-04-21	
2	PRIOR APPLICATION NUMBER: 60/082704	
3	PRIOR FILING DATE: 1998-04-22	
4	PRIOR APPLICATION NUMBER: 60/082797	
5	PRIOR FILING DATE: 1998-04-22	
6	PRIOR APPLICATION NUMBER: 60/083322	
7	PRIOR FILING DATE: 1998-04-28	
8	PRIOR APPLICATION NUMBER: 60/083495	
9	PRIOR FILING DATE: 1998-04-29	
10	PRIOR APPLICATION NUMBER: 60/083496	
11	PRIOR FILING DATE: 1998-04-29	
12	PRIOR APPLICATION NUMBER: 60/083499	
13	PRIOR FILING DATE: 1998-04-29	
14	PRIOR APPLICATION NUMBER: 60/083559	
15	PRIOR FILING DATE: 1998-04-29	
16	PRIOR APPLICATION NUMBER: 60/083466	
17	PRIOR FILING DATE: 1998-05-05	
18	PRIOR APPLICATION NUMBER: 60/084414	
19	PRIOR FILING DATE: 1998-05-06	
20	PRIOR APPLICATION NUMBER: 60/084639	
21	PRIOR FILING DATE: 1998-05-07	
22	PRIOR APPLICATION NUMBER: 60/084640	
23	PRIOR FILING DATE: 1998-05-07	
24	PRIOR APPLICATION NUMBER: 60/084643	
25	PRIOR FILING DATE: 1998-05-07	
26	PRIOR APPLICATION NUMBER: 60/085573	
27	PRIOR FILING DATE: 1998-05-15	
28	PRIOR APPLICATION NUMBER: 60/085579	
29	PRIOR FILING DATE: 1998-05-15	
30	PRIOR APPLICATION NUMBER: 60/085580	
31	PRIOR FILING DATE: 1998-05-15	
32	PRIOR APPLICATION NUMBER: 60/085582	
33	PRIOR FILING DATE: 1998-05-15	
34	PRIOR APPLICATION NUMBER: 60/085700	
35	PRIOR FILING DATE: 1998-05-15	
36	PRIOR APPLICATION NUMBER: 60/086023	
37	PRIOR FILING DATE: 1998-05-18	
38	PRIOR APPLICATION NUMBER: 60/086392	
39	PRIOR FILING DATE: 1998-05-22	
40	PRIOR APPLICATION NUMBER: 60/086486	
41	PRIOR FILING DATE: 1998-05-22	
42	PRIOR APPLICATION NUMBER: 60/087098	
43	PRIOR FILING DATE: 1998-05-28	
44	PRIOR APPLICATION NUMBER: 60/087208	
45	PRIOR FILING DATE: 1998-05-28	
46	PRIOR APPLICATION NUMBER: 60/087609	
47	PRIOR FILING DATE: 1998-06-02	
48	PRIOR APPLICATION NUMBER: 60/087759	
49	PRIOR FILING DATE: 1998-06-02	
50	PRIOR APPLICATION NUMBER: 60/087827	
51	PRIOR FILING DATE: 1998-06-03	
52	PRIOR APPLICATION NUMBER: 60/088025	
53	PRIOR FILING DATE: 1998-06-04	
54	PRIOR APPLICATION NUMBER: 60/088028	
55	PRIOR FILING DATE: 1998-06-04	
56	PRIOR APPLICATION NUMBER: 60/088029	
57	PRIOR FILING DATE: 1998-06-04	
58	PRIOR APPLICATION NUMBER: 60/088033	
59	PRIOR FILING DATE: 1998-06-04	
60	PRIOR APPLICATION NUMBER: 60/088167	
61	PRIOR FILING DATE: 1998-06-05	
62	PRIOR APPLICATION NUMBER: 60/088202	
63	PRIOR FILING DATE: 1998-06-05	
64	PRIOR APPLICATION NUMBER: 60/088212	
65	PRIOR FILING DATE: 1998-06-05	
66	PRIOR APPLICATION NUMBER: 60/088217	
67	PRIOR FILING DATE: 1998-06-05	
68	PRIOR APPLICATION NUMBER: 60/088326	
69	PRIOR FILING DATE: 1998-06-04	
70	PRIOR APPLICATION NUMBER: 60/088655	
71	PRIOR FILING DATE: 1998-06-09	
72	PRIOR APPLICATION NUMBER: 60/088722	
73	PRIOR FILING DATE: 1998-06-10	

PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088740
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088811
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088825
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088863
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089090
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653

Query Match 6.3%; Score 180; DB 4; Length 596;
Best Local Similarity 19.2%; Pred. No. 0.0031;

Matches 109; Conservative 90; Mismatches 272; Indels 98; Gaps 19;

4 GNLGNPNVNNISIPAPPLPSOTDAGGRGQLINSTGPLGRALFPVRNSMADSGDNRA 63
13 GLLHLEAATNS-----NETSTANTGSSVTSISG-----ASTATNG 49
64 SDVPGLPVNPMLAASEITLNDGFEVLHDGPLDTLNROIGSSV-PRVETEDGKHIAVG 122
50 SSVTSSGCVTATISGSSVTSN-GVSIIV-TNSEPHITSSGISTATNSEFTASGIGIATN 107
123 QNRGVETS--VVLSDOEVARLOSIDPEGRKDFVTCGRGAGHAMVTVASDITEARQRL 180
108 SSSSTTSSGASTATNSE---SSTPSSGASTVTNSSSVTSSGASTATNSESTVSSRAS 163
181 ELLEPKGTGESKAGSKGVGELRESNGAENTTETOTSTSSLSLSDPKMLALCTVAT 240
164 TATNSESTTSSGASTATN-SDSSTTSSGASTATNSESTTSSG-----ASTAT 211
241 GLIGLAATGIVQALALTPPEPDSPTTDPDAAASATETATRDOLTEKAFQNPONOKYNIDE 300
212 N-----SESTVSSRASTATNSESTT---SSGASTATNSESRITSSGAGATATNSESTTS 264
301 LG-----NAIPSGVLKDVVANIEEOAKAAGEAKQAIEENNAQOKKYDEQOAKR 351
265 SGASTATNDSSTVSSGA---STATNSESTTSSGAST-----ATN 302
352 OEELKVSAGAGYGL-SCALILGGGIGVAVTAAALHRKNQVPEQTTTTTTTTTTSARTVEN 410
303 SSSSTTSSGASTATNDSSTTSSGAGATATNSESTVSSGISTVTNSESSTPSSGANTATN 362
411 KPANTPAGNVDTPESEDPTMESRRSMASTSTFPDTSIGTVQNPVADVTS-----464
363 SSSSTTSSGANTAT-NSESTVSSGASTATNSESTTSSGAGVSTATNSESTTSSGASTAT 421
465 LHDQVPTNSNTSVONKMGNTDSVVYSTIOHPPRDTTNGARLLGNPSAGIOSTYARLAL 524
422 NDSSTTSSSEASTATATNSESTVSSGISTVTNSESSTTSSGANTATNCGSVTS-----474
525 SGGLRHDMGGLTGGSNSAVNTSNPPAPG 553

Db 475 AGSGTAAALGCMHTTSHS-STAVSEAKPG 502

RESULT 15
US-10-176-483-310
Sequence 310, Application US/10176483
Publication No. US20030017541A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C68
CURRENT APPLICATION NUMBER: US/10/176,483
CURRENT FILING DATE: 2002-06-20
Prior application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 310
LENGTH: 596
TYPE: PRT
ORGANISM: Homo Sapien
US-10-176-483-310

Query Match 6.3%; Score 180; DB 4; Length 596;
Best Local Similarity 19.2%; Pred. No. 0.0031;

Matches 109; Conservative 90; Mismatches 272; Indels 98; Gaps 19;

4 GNLGNPNVNNISIPAPPLPSOTDAGGRGQLINSTGPLGRALFPVRNSMADSGDNRA 63
13 GLLHLEAATNS-----NETSTANTGSSVTSISG-----ASTATNG 49
64 SDVPGLPVNPMLAASEITLNDGFEVLHDGPLDTLNROIGSSV-PRVETEDGKHIAVG 122
50 SSVTSSGCVTATISGSSVTSN-GVSIIV-TNSEPHITSSGISTATNSEFTASGIGIATN 107
123 QNRGVETS--VVLSDOEVARLOSIDPEGRKDFVTCGRGAGHAMVTVASDITEARQRL 180
108 SSSSTTSSGASTATNSE---SSTPSSGASTVTNSSSVTSSGASTATNSESTVSSRAS 163
181 ELLEPKGTGESKAGSKGVGELRESNGAENTTETOTSTSSLSLSDPKMLALCTVAT 240
164 TATNSESTTSSGASTATN-SDSSTTSSGASTATNSESTTSSG-----ASTAT 211
241 GLIGLAATGIVQALALTPPEPDSPTTDPDAAASATETATRDOLTEKAFQNPONOKYNIDE 300
212 N-----SESTVSSRASTATNSESTT---SSGASTATNSESRITSSGAGATATNSESTTS 264
301 LG-----NAIPSGVLKDVVANIEEOAKAAGEAKQAIEENNAQOKKYDEQOAKR 351
265 SGASTATNDSSTVSSGA---STATNSESTTSSGAST-----ATN 302
352 OEELKVSAGAGYGL-SCALILGGGIGVAVTAAALHRKNQVPEQTTTTTTTTTTSARTVEN 410
303 SSSSTTSSGASTATNDSSTTSSGAGATATNSESTVSSGISTVTNSESSTPSSGANTATN 362
411 KPANTPAGNVDTPESEDPTMESRRSMASTSTFPDTSIGTVQNPVADVTS-----464
363 SSSSTTSSGANTAT-NSESTVSSGASTATNSESTTSSGAGVSTATNSESTTSSGASTAT 421
465 LHDQVPTNSNTSVONKMGNTDSVVYSTIOHPPRDTTNGARLLGNPSAGIOSTYARLAL 524
422 NDSSTTSSSEASTATATNSESTVSSGISTVTNSESSTTSSGANTATNCGSVTS-----474

QY 525 SGGLRHDMGGLTGGSNSAVTSTNNPPAPG 553
 475 AGSGTAALTGMHTTSHA-STAVSEAKPG 502
 Db

Search completed: August 1, 2006, 22:29:00
Job time : 186 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 1, 2006, 22:26:05 ; Search time 32 Seconds
(without alignments)
1151.645 Million cell updates/sec

Title: US-09-189-415D-11

Perfect score: 2840

Sequence: 1 MPKINGLGHNPVNNNSIPAP.....SNSAVNTSNNPAPGSHRFV 558

Scoring table: BLOSUM62

Searched: Gapop 10.0 , Gapext 0.5

Total number of hits satisfying chosen parameters: 232337

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA New.*
1: /EMC_Celerra_SIDS3/Ptodata/2/pubpaa/US09_NEW_PUB pep.*
2: /EMC_Celerra_SIDS3/Ptodata/2/pubpaa/US06_NEW_PUB pep.*
3: /EMC_Celerra_SIDS3/Ptodata/2/pubpaa/US07_NEW_PUB pep.*
4: /EMC_Celerra_SIDS3/Ptodata/2/pubpaa/US08_NEW_PUB pep.*
5: /EMC_Celerra_SIDS3/Ptodata/2/pubpaa/PCT_NEW_PUB pep.*
6: /EMC_Celerra_SIDS3/Ptodata/2/pubpaa/US11_NEW_PUB pep.*
7: /EMC_Celerra_SIDS3/Ptodata/2/pubpaa/US11_NEW_PUB pep.*
8: /EMC_Celerra_SIDS3/Ptodata/2/pubpaa/US60_NEW_PUB pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	180	6.3	596	US-10-196-749-310	Sequence 310, App
2	180	6.3	596	US-11-101-316-100	Sequence 100, App
3	180	6.3	596	US-11-376-673-100	Sequence 100, App
4	180	6.3	1086	US-11-330-403-3535	Sequence 3535, App
5	169	6.0	2271	US-10-471-571A-3948	Sequence 3948, App
6	168.5	5.9	2804	US-11-330-403-12380	Sequence 12380, App
7	160	5.6	22152	US-10-544-944-1	Sequence 1, App1
8	151	5.3	1306	US-11-251-208-6	Sequence 6, App1
9	149.5	5.3	785	US-11-317-330A-19	Sequence 19, App1
10	149	5.2	2478	US-10-471-571A-2278	Sequence 2278, App
11	147	5.2	985	US-11-342-171-6	Sequence 6, App1
12	147	5.2	985	US-11-342-171-66	Sequence 66, App1
13	145.5	5.1	2053	US-11-013-711-9	Sequence 9, App1
14	145	5.1	2314	US-11-013-711-11	Sequence 11, App1
15	143	5.0	804	US-11-056-355B-87568	Sequence 87568, App
16	143	5.0	915	US-10-523-014-4	Sequence 4, App1
17	143	5.0	935	US-10-449-902-42274	Sequence 42274, App
18	141	5.0	1071	US-11-056-355B-80594	Sequence 80594, App
19	140.5	4.9	1316	US-11-121-154-83	Sequence 83, App1
20	138	4.9	788	US-10-449-902-41167	Sequence 41167, App
21	137	4.8	1501	US-10-630-629-4	Sequence 4, App1
22	137	4.8	1992	US-11-013-711-3	Sequence 3, App1
23	137	4.8	1992	US-11-013-711-13	Sequence 13, App1
24	137	4.8	2047	US-11-013-711-4	Sequence 4, App1
25	137	4.8	2047	US-11-013-711-7	Sequence 7, App1

26	136.5	4.8	1498	US-11-266-446-94	Sequence 94, App1
27	136	4.8	791	US-11-121-154-24	Sequence 24, App1
28	135.5	4.8	1581	US-11-283-329-162	Sequence 162, App
29	134	4.7	288	US-11-321-421-341	Sequence 341, App
30	134	4.7	288	US-11-321-421-344	Sequence 344, App
31	133.5	4.7	1152	US-10-449-902-41197	Sequence 41197, App
32	133.5	4.7	1531	US-11-333-747A-44	Sequence 44, App1
33	132.5	4.7	849	US-10-449-902-43102	Sequence 43102, App
34	132	4.6	2659	US-11-256-173-28	Sequence 28, App1
35	132	4.6	9535	US-10-471-571A-4496	Sequence 4496, App
36	131.5	4.6	549	US-10-449-902-54423	Sequence 54423, App
37	131.5	4.6	710	US-10-449-902-41319	Sequence 41319, App
38	131	4.6	1025	US-10-449-902-42266	Sequence 42266, App
39	130.5	4.6	3429	US-10-953-349-9275	Sequence 9275, App
40	130.5	4.6	3438	US-10-953-349-9274	Sequence 9274, App
41	130.5	4.6	3478	US-10-953-349-9273	Sequence 9273, App
42	130	4.6	1303	US-10-449-902-55274	Sequence 55274, App
43	130	4.6	7362	US-11-330-403-9773	Sequence 9773, App
44	128.5	4.5	799	US-10-449-902-50884	Sequence 50884, App
45	128	4.5	365	US-11-056-355B-81599	Sequence 81599, App

ALIGNMENTS

RESULT 1
US-10-196-749-310
Sequence 310, Application US/10196749
GENERAL INFORMATION:
Publication No. US20060094864A1
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
TITLE OF INVENTION: ACIDS ENCODING THE SAME
FILE REFERENCE: P3430R1C340
CURRENT APPLICATION NUMBER: US/10/196,749
CURRENT FILING DATE: 2002-07-16
PRIOR APPLICATION NUMBER: 10/052586
PRIOR FILING DATE: 2002-01-15
PRIOR APPLICATION NUMBER: 60/059263
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/059266
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/063120
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063121
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063486
PRIOR FILING DATE: 1997-10-21
PRIOR APPLICATION NUMBER: 60/063540
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063541
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063544
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 310
LENGTH: 596
TYPE: PRT
ORGANISM: Homo Sapien
US-10-196-749-310

```
Query Match          6.3%: Score 180; DB 6; Length 596;
Best Local Similarity 19.2%: Pred. No. 0.0066;
Matches 109; Conservative 90; Mismatches 272; Indels 98; Gaps 19;

QY 4 GNLGHPNVNNSIPAPPLPSQTDGAGRGQLINSTGELGRALFTPVNSMADSGDNRA 63
DB 13 GLLHLLEBATNS-----NETSTSNANTGSSVSISSG-----ASTATNSG 49
QY 64 SDVPEGLPVNPKRLAASEITLNDGEVLHDHGLDPLTLNROISSV-FRVEIOEDGKHIVG 122
DB 50 SSVTSSGVSTATIGSSSVTSN-GVSIIV-TNSEFHITSSGISTATNSESTSSGSIATN 107
QY 123 QNRGVETS--VVLSDOEYARLQSIDPEGKDKFVFTGRCGAGHAMVTVASDITEARORIL 180
DB 108 SESSTSSGASTATNSE-----STPSSGASTVNTSSGSSVTSNGASTATNSESTVSSRAS 163
QY 181 ELLEPKGTGSKGAGESKGVGELRESNGAENTTETGTSTSSLRSDPKMLALGVAT 240
DB 164 TATNSESTSSGASTATN-SDSSTSSGASTATNSESTSSG-----ASTAT 211
QY 241 GLIGLAATGYQALALTPEDPSPTTDDPAASAATETATROULTKEAFQNPNOKNINDE 300
DB 212 N-----SESTVSSRASTATNSESTT---SSGASTATNSESTTSSNGAGTATNSESTTS 264
QY 301 LG-----NAIPSGVLKODVVANIEQAKAGEAKQQAENNAQAOKKYDEQAKR 351
DB 265 SGASTATNSDSSTVSSGA---STATNSESTSSGAST-----ATN 302
QY 352 QEELKVSAGAGYGL-SGALLIGGIGVAVTALHRKQNPVEQTTTTTTTTTSARYEN 410
DB 303 SESSTSSGASTATNSDSSTSSGAGTATNSESTVSSGISTVNTSESTSSGANTATN 362
QY 411 KPANTPQAQNVDTGSEPTMESRRSSMASTSTFFDTSIGTYONPADVXTS-----464
DB 363 SESSTSSGANTAT-NSSESTVSSGASTATNSESTSSGVTATNSESTSSGASTAT 421
QY 465 LHDGVPNTSNSTSVQNNKNTDSVYVSTIOHPPTDTNGARLLGNPAGIOSTYARIAL 524
DB 422 NSDSSTSSSEASTATNSESTVSSGISTVNTSESTSSGANTATNCGSSVTS-----474
QY 525 SGGLRHDMGGLTGGSNSAVNTSNPPAPG 553
DB 475 AGSGTAALTGKHTTSSHA-STAVSEAKPG 502

RESULT 2
US-11-101-316-100
; Sequence 100, Application US/11101316
; Publication No. US2006009657A1
; GENERAL INFORMATION:
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: ANTIBODIES TO A POLYPEPTIDE ENCODED BY A NUCLEIC ACID
; FILE REFERENCE: P323ORIC17C1
; CURRENT APPLICATION NUMBER: US/11/101,316
; PRIOR FILING DATE: 2005-04-06
; PRIOR APPLICATION NUMBER: 10/063526
; PRIOR FILING DATE: 2002-05-03
; PRIOR APPLICATION NUMBER: 10/006867
; PRIOR FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: PCT/US00/23328
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 09/380137
; PRIOR FILING DATE: 1999-08-25
; PRIOR APPLICATION NUMBER: PCT/US99/12252
; PRIOR FILING DATE: 1999-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
```

```
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 100
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Homo Sapien
US-11-101-316-100

Query Match          6.3%: Score 180; DB 7; Length 596;
Best Local Similarity 19.2%: Pred. No. 0.0066;
Matches 109; Conservative 90; Mismatches 272; Indels 98; Gaps 19;

QY 4 GNLGHPNVNNSIPAPPLPSQTDGAGRGQLINSTGELGRALFTPVNSMADSGDNRA 63
DB 13 GLLHLLEBATNS-----NETSTSNANTGSSVSISSG-----ASTATNSG 49
QY 64 SDVPEGLPVNPKRLAASEITLNDGEVLHDHGLDPLTLNROISSV-FRVEIOEDGKHIVG 122
DB 50 SSVTSSGVSTATIGSSSVTSN-GVSIIV-TNSEFHITSSGISTATNSESTSSGSIATN 107
QY 123 QNRGVETS--VVLSDOEYARLQSIDPEGKDKFVFTGRCGAGHAMVTVASDITEARORIL 180
DB 108 SESSTSSGASTATNSE-----STPSSGASTVNTSSGSSVTSNGASTATNSESTVSSRAS 163
QY 181 ELLEPKGTGSKGAGESKGVGELRESNGAENTTETGTSTSSLRSDPKMLALGVAT 240
DB 164 TATNSESTSSGASTATN-SDSSTSSGASTATNSESTSSG-----ASTAT 211
QY 241 GLIGLAATGYQALALTPEDPSPTTDDPAASAATETATROULTKEAFQNPNOKNINDE 300
DB 212 N-----SESTVSSRASTATNSESTT---SSGASTATNSESTTSSNGAGTATNSESTTS 264
QY 301 LG-----NAIPSGVLKODVVANIEQAKAGEAKQQAENNAQAOKKYDEQAKR 351
DB 265 SGASTATNSDSSTVSSGA---STATNSESTSSGAST-----ATN 302
QY 352 QEELKVSAGAGYGL-SGALLIGGIGVAVTALHRKQNPVEQTTTTTTTTTSARYEN 410
DB 303 SESSTSSGASTATNSDSSTSSGAGTATNSESTVSSGISTVNTSESTSSGANTATN 362
QY 411 KPANTPQAQNVDTGSEPTMESRRSSMASTSTFFDTSIGTYONPADVXTS-----464
DB 363 SESSTSSGANTAT-NSSESTVSSGASTATNSESTSSGVTATNSESTSSGASTAT 421
QY 465 LHDGVPNTSNSTSVQNNKNTDSVYVSTIOHPPTDTNGARLLGNPAGIOSTYARIAL 524
DB 422 NSDSSTSSSEASTATNSESTVSSGISTVNTSESTSSGANTATNCGSSVTS-----474
QY 525 SGGLRHDMGGLTGGSNSAVNTSNPPAPG 553
DB 475 AGSGTAALTGKHTTSSHA-STAVSEAKPG 502

RESULT 3
US-11-376-673-100
; Sequence 100, Application US/11376673
; Publication No. US20060160186A1
; GENERAL INFORMATION:
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: A NUCLEIC ACID UNDEREXPRESSED IN STOMACH TUMOR AND
; FILE REFERENCE: P323ORIC165C
; CURRENT APPLICATION NUMBER: US/11/376,673
; PRIOR FILING DATE: 2006-03-14
; PRIOR APPLICATION NUMBER: 10/063742
; PRIOR FILING DATE: 2002-05-09
; PRIOR APPLICATION NUMBER: 10/006867
; PRIOR FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: PCT/US00/23328
; PRIOR FILING DATE: 2000-08-24
```

: PRIOR APPLICATION NUMBER: 60/170262
 : PRIOR FILING DATE: 1999-12-09
 : NUMBER OF SEQ ID NOS: 170
 : SEQ ID NO 100
 : LENGTH: 596
 : TYPE: PRT
 : ORGANISM: Homo Sapien
 US-11-376-673-100

Query Match	6.3%	Score 180;	DB 7;	Length 596;
Best Local Similarity	19.2%;	Pred. No. 0.0066;		
Matches 109;	Conservative 90;	Mismatches 272;	Indels 98;	Gaps 19;

```

QY 4 GNLGHPNNVNS.IPPAPPLP$QTDGAGGGGQOLIN$TGPELGSALFTTPVRNMAOSGNRA 63
Db 13 GLLHLLEATNS-----NET$T$ANTGSSVIGS-----ASTATNSG 49
QY 64 SDVPGLEPVNPMRLA$EITLNDGFEVYLDHGLDPLTNROIGSV-FVETQEDKHI$AVG 122
Db 50 $SYTSSGCV$TATIGSS$VTSN-GV$IV-TNSEFHTTSSGISTATNSEP$T$ASG$ISIATN 1070
QY 123 Q$R$GVETS--V$LSDEYARLOSIDPEGDKRVFTG$GAGG$HAMVYASDIT$ARQ$IL 160
Db 108 $E$STTSSG$ASTATN$E---$T$P$SG$ASTVYTN$GSS$VTS$G$ASTATN$E$STV$SR$S 163
QY 181 ELLEP$KGT$ESK$AG$E$K$G$V$G$EL$RES$N$G$AEVTE$T$OT$T$S$T$SL$R$D$PK$ML$G$V$AT 240
Db 164 TATN$E$STT$SS$AS$AT$ATN-$D$S$STT$SS$AS$AT$ATN$E$STT$S$G-----$A$T$AT 2111
QY 241 GLIGL$AT$GI$VQ$AL$T$P$E$D$P$T$T$T$D$P$A$A$S$AT$E$T$ARD$Q$LEAFQ$ND$K$V$N$IDE 3000
Db 212 N-----S$E$T$V$S$R$A$T$ATN$E$S$S$T$T$--$S$G$A$T$ATN$E$S$R$T$T$N$G$A$T$ATN$E$S$T$T$S 264
QY 301 LG-----NAI$P$G$V$L$K$D$V$V$AN$IE$Q$A$A$A$E$E$K$Q$A$I$E$N$N$A$Q$K$D$E$Q$A$K$R 351
Db 265 S$G$A$T$ATN$D$S$T$V$S$G$A-----$T$ATN$E$S$T$T$S$G$A$T-----ATN 3020
QY 352 Q$E$L$K$V$S$G$G$Y$G$L-$S$G$A$L$IG$G$G$V$N$T$A$L$H$R$K$N$Q$P$V$O$T$T$T$T$T$T$T$T$T$S$A$T$V$E$N 4110
Db 303 S$E$S$T$T$S$G$A$T$ATN$D$S$T$T$S$G$A$T$ATN$E$S$T$V$S$G$I$T$V$N$E$S$T$T$S$G$A$N$AT$ATN 3620
QY 411 KP$A$N$T$P$A$Q$N$V$D$T$P$E$S$E$D$T$E$M$E$R$R$S$M$A$S$T$E$T$F$D$S$IG$V$O$N$P$A$D$V$K$T$S----- 4644
Db 363 S$E$S$T$T$S$G$A$N$AT-$N$E$S$S$T$V$S$G$A$T$ATN$E$S$T$T$S$G$V$T$ATN$E$S$T$T$S$G$A$T$AT 4211
QY 465 LHD$Q$V$T$N$S$N$T$S$V$O$N$M$G$N$D$S$V$V$Y$S$T$IQ$H$P$R$D$T$N$G$A$R$L$GN$P$AG$I$Q$S$T$Y$A$R$L$AL 5244
Db 422 N$D$S$T$T$S$E$A$T$AT$N$E$S$S$T$V$S$G$I$F$V$T$N$E$S$T$T$S$G$A$N$T$ATN$G$S$S$V$T$S----- 4744
QY 525 S$G$L$R$H$D$M$G$L$T$G$G$S$A$N$T$S$N$N$P$A$G 553
Db 475 A$G$G$T$A$L$T$G$M$T$T$H$S$A-$T$A$V$E$A$K$Q 502

```

Matches 111; Conservative 78; Mismatches 243; Indels 152; Gaps 20;

```

0Y 3 IGNLGHN-PNVNNSI:PPA-----PPLPEQOTGAGRGQILINSTGLGSAALFTPVNSM 55
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 314 LGSVAYNPGVPNNNAVPASNCCKYVPSG-TD-----BSSSLST-----ST 354
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
0Y 56 ADSGNRASDVPGLEVPNPRLAASEITLNDGFEVLHDHGPLDTLNRQIGSSVFRVETOED 115
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 355 SEQSSSSATS-V-----SASETS-----DTSSESSSSSEVSSTQE 390
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
0Y 116 GKHIIVGQRNGVETSVLSDQFAYALQSIDPEGKCKFVETGGRGAGHAMVYASDTEA 175
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 391 PSSSTPESSSSETS--STOESSSTECPSSSDTSSTEASSSESSSTAPSSTEAASSTES 447
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
0Y 176 RQRIELLEPKGTGSKAGESKYGELRESNGAENT-ETQTSSTSLRSDPKMLA 234
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 448 STE-----EPSSSTE-----GSSSESSSESSSTQEPSSSTKESSSTGCP----- 489
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
0Y 235 LGVATVGLIGLAATGIVQALALTPPEPSFTPTDPPAASATETATRDQLTKEAFQNPNO 294
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 490 -----STESSSTEGPSSST-----DSDTITASSTDEQSSSGTGQSTE 531
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
0Y 295 KVNID---ELGNALPSGVLKDDVYANIEBQAKAAGEEKQOALINNAQAQKYDEQOAK 350
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 532 DEPIDSTESDTSATDSSTATDSATNTDNTSESDTSSTATDTSSTDSSTASSTENTDV 591
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
0Y 351 RQELKVESGAGYGLSALLIGGIGVANTALHRKNQVEQTTTTTTTTTTSARTVEN 410
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 592 TDSSTDSVTGA-----TESSTATDNTDADSSSTVE 623
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
0Y 411 KPANNTPAQGVNDTPGESDTPMESRRSRMASTSTFPTSSIGTVQNPYADVKTSLHSDQV 470
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 624 TGATDTSSTATDNTNGATESSDTSNNGATESSTATDNTSATNTDNTNGSNTAINT--DDNT 682
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
0Y 471 PTSNSNTSVQNMGNDSVVYST-----IQHPDRTDND-GARLLANPSAGIOSYARLA 523
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 683 ATDTSSTETNTAINTDGETEINTGATEINTDTSASNTDMDWTGNSNTAINTGTDINTDNTN-- 740
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
0Y 524 LSGGRHDMGGLTGS-----GNSAVNTSN 548
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 741 -TGGDNTNGTNTGGTDTKTGNTATGANTGANTGATETNTAINTNGN 783
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

Query Match	6.3%;	Score 180;	DB 7;	Length 1086;
Best Local Similarity	19.0%;	Pred. No. 0.014;		

```
QY      24 SQTGAGGGRGLNSTGTPLGSRALFTVVRNSMADS-----GDNRASDVPGLPVN 72
```

```
Db 1620 SRSTASSTSMRTSTSDSQMSLSTSTSTMSDSTSLSDSVSDSTSDSTASSTSGMSV 1679
Qy 73 PMRLASEITLNDGEVLHDHGPLDTLNRQIGSSVFVETQEDGKHAIVGONGVETSV 132
Db 1680 SISLSDSTSTSTASV-----MSASISDSQSMESVNDSE--SVSESNSSDSKS 1728
Qy 133 LSDGEYARLQSIDPEGKDFVFTGGGAGHAMVTVAADITEARORILELPEKGTGSK 192
Db 1729 MSGSTSV-----SDSGSLSVSTSLRKSRSVSESSSLSCSQMSDSV 1769
Qy 193 GAGESKGVG---ELRESNSGAEN-----TTTGTSTSTSLRSDPKLMLALGVATGL 242
Db 1770 STSDSVSLSVSTSLRSESSESDSLSDSKSTSGSTSTSGSLS-----TSTSL 1819
Qy 243 IG---LMTGIVQALALTPEDSPPTTDPD-----AAASATERATRDQTLKEAFQPD 292
Db 1820 SGSESVSTSLSDSISMS--DSTSTSDSDSLSGSISLISGSLSTSDSDS----- 1870
Qy 293 NQKVNIDELGNALPESVLKDDVYANIEQAKAAGEAKQQA1--ENNAQAQKXYDEQA 349
Db 1871 -----KSLSSSQSMGSESTSTSVSDSQSSSTSGSDSMI 1907
Qy 350 KRQELKYSAGAGLGLGAL-----ILGGIGVAVTALHRKNQVPEQTT---T 395
Db 1908 SASBEDSMSTSDSSISGSNSTSTSLSTSDSMGSGSVSTSTSL--SDSISGSTSVSDS 1964
Qy 396 TTTTTSARTVKKPANNTPAOGNVDTP-----PGSEDTMESRRSSMA--STSTTF 446
Db 1965 SSTSTSTSLSDMSQSGSTSTASGSLSTSTSTSMGASSTSSGSTSVSTSTSDSIS 2024
Qy 447 DTSSI-----GVQNPYADVKTSLHDSQ-VPTSNSTSVQNMGTDSVVYSTIOHPPRD 499
Db 2025 DSTSISISGSGSTVSESTSDSTSLSDSESLSTSDSDSTSTSTSGSTSTSTSESL 2084
Qy 500 TTDNGCARLLGNPSAGIQSTYARLALSGGRHDMGELTGGSNS-AVNTSN 548
Db 2085 TSGSGSTSVSDSTSMSESNSSSVMS---QDKSDSTSLSDSESVSTST 2130

RESULT 6
US-11-330-403-12380
; Sequence 12380, Application US/11330403
; Publication No. US20060159563A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53629)B
; CURRENT APPLICATION NUMBER: US/11/330,403
; CURRENT FILING DATE: 2006-01-12
; NUMBER OF SEQ ID NOS: 19250
; SEQ ID NO 12380
; LENGTH: 2804
; TYPE: PRT
; ORGANISM: Magnetococcus sp. MC-1
US-11-330-403-12380

Query Match 5.9%; Score 168.5; DB 7; Length 2804;
Best Local Similarity 21.9%; Pred. No. 0.21;
Matches 135; Conservative 70; Mismatches 224; Indels 187; Gaps 25;

Qy 20 PPLPQGTGAGGAGGQLINSTGPLGRALFTPVRSMA--DSGNRRASDVYGLVNPMLRAA 78
Db 698 PSTVASVTRAGGHDVYINASDGLSLGWSAVGSVTTIDGCGSLDNNGATANIANS 757
Qy 79 S-----ETLNDGEVLHDHGPLDTLNRQIGSSVFVETQEDGKHAIVGONGVETSV 132
Db 758 SIKSDKAGTASDALET-----TVDTL-----AVEITSGKSFYIDESDA----- 797
Qy 133 LSDGEYARLQSIDPE--GKDKFVTGGGAGHAMVTVAADITEARORILELPEKGTG 189
Db 798 -----LTSINAVNNGSTNLNFTG-----SFAFNAT----- 824
Qy 190 BSKGAGESKGVGELAESNG-----AENTTGTSTSTSLRSDPKLMLAL 235
```

```
Db 825 --TAFSSTGVSVTFENTGGVAIGVTATGSGATITASTATDATSAITADTVV-LTA 881
Qy 236 GT-----VATGLIGLAATGIV-----QALALTPEDSPPTTDP----- 268
Db 882 GTSIGASGSTTKSTRATTLTYLASAGSTYAOESDGAATVNAKAVGCKNIERATTTGNLTIG 941
Qy 269 --DAASATETATPDQTLKEAFQNDQKVNIDELGNAL-PSGVLKDDVYANIEQAKAA 325
Db 942 TISAKGSYTLTATAGSVLSSGTSATGATVVLASGATLIGASGASVMTVAANLATATASAG 1001
Qy 326 GEEAKQQA1ENNAQAQKXYDEQAKROBELKVSSGAGYGLS--GALLIGGIGVAVTAL 383
Db 1002 G-----VYLSNAGDVTLLTAVATGAGFQLSNTGALVING-----VPA 1039
Qy 384 HRKNQVPEQTTTTTTTTTSARTVKKPANNTPAQ-----GVNDPGSRDPTMESR 434
Db 1040 --QQA1-----SLTASGATIDGNGATNNISAESLTLVGLSIGSTATGAVDTP----- 1084
Qy 435 RSSMASTSTPEFDTSSICTGVQNPYADVKTSLHDSQVPTSNSTSVQNMGTDSVVS--- 491
Db 1085 -----QVNSVTRATTTSGGIYINELSGGLTITATVQSDANVSLTGAGDIALGVITAKG 1139
Qy 492 --TIQHPPRDTTNGA-----RLG-NPSAGIQSTYARLAL-----SGGIRH 530
Db 1140 DDTVLSAKKITDNNGANNTADILNVTGNGVEGLESTITQLSTSGSTDIVANAGAMAI 1189
Qy 531 DMGELTGGSNAVNTS 546
Db 1200 TKASLEGGSSSFFIAES 1215

RESULT 7
US-10-544-944-1
; Sequence 1, Application US/10544944
; Publication No. US20060134120A1
; GENERAL INFORMATION:
; APPLICANT: Diamandis, Eleftherios P.
; TITLE OF INVENTION: Multiple Marker Assay for Detection of Ovarian Cancer
; FILE REFERENCE: 11757.0088USMO
; CURRENT APPLICATION NUMBER: US/10/544,944
; CURRENT FILING DATE: 2005-08-09
; PRIOR APPLICATION NUMBER: PCT/CA2004/000281
; PRIOR FILING DATE: 2004-02-26
; PRIOR APPLICATION NUMBER: US 60/450,406
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 22152
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (13877)..(13878)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (13880)..(13880)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (13887)..(13887)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (13890)..(13891)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (13893)..(13893)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
; FEATURE:
```


RESULT 9

US-11-317-330A-19
Sequence 19, Application US/11317330A
Publication No. US20060154240A1
GENERAL INFORMATION:
APPLICANT: MCCRAY, JR, PAUL B.
APPLICANT: SINN, PATRICK L.
APPLICANT: VOYTAS, DAN
APPLICANT: DAI, JUNBIAO
TITLE OF INVENTION: COMPOSITIONS AND METHODS RELATED TO MODIFIED RETROVIRAL
TITLE OF INVENTION: VECTORS FOR RESTRICTED, SITE SPECIFIC INTEGRATION
FILE REFERENCE: IOMA:074US
CURRENT APPLICATION NUMBER: US/11/317,330A
CURRENT FILING DATE: 2005-12-22
PRIOR APPLICATION NUMBER: 60/638,590
PRIOR FILING DATE: 2004-12-22
NUMBER OF SEQ ID NOS: 27
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 19
LENGTH: 785
TYPE: PRF
ORGANISM: Homo sapiens
US-11-317-330A-19

Query Match 5.3%; Score 149.5; DB 7; Length 785;
Best Local Similarity 20.6%; Pred. No. 0.5; 260; Indels 141; Gaps 27;
Matches 127; Conservative 88; Mismatches 141; Gaps 27;

QY 4 GNLGHNPNVNSIPPAAPLPSSQTDGAGRGOLINSTGPGSRA-----LFTPVNSMADS 58
DB 23 GNNNGNGNGGAFSQARSSSTGSSSTGGGSGESQSPALALATCSRIESPNSNNNG 82
QY 59 GGNRASDVGLFVNPMRLAASEITLNDGEVL-HDHGPDLTNRQIGSSVFRVETQEDGK 117
DB 83 GPGSGGTGELDLTATQLSQGA-----NGWQIISSSSGATPTEKSGSSTNGSGESSK 138
QY 118 HIAVG-----ORNGVETSY--VLSDQEV--ARLOSIDEGDKVFVTCG--- 157
DB 139 NRTVGCGQYVVAAPLQOQVLTGPGWPMNIQOVIPQFQTVDQ-OLQFPAATGAQVO 197
QY 158 RGAGHAMVTVAADITEARQRIELELEPKGTGSKGAGESKVGELRESNSGAENTTETQ 217
DB 198 QDGSQIQI-----IPGANQOI---ITNRSGSGNIIAAMPNLLQOQVPLQGLANNVLSQ 249
QY 218 TSTSTSSLRSPDKMLAGTVAATGLGLAATGIVQALALTPPEPDSPTTDPAAASATET 277
DB 250 TQYVYN-----VVALNGNITLLPVNSVSAATLTPSSQAVTTISSGSGSGSQP 298
QY 278 ATRDQLTKEA-----FQNPDNK-----VVIDELGNA-----IP-- 306
DB 299 VTSQTTISSASLVSSQASSSFTTANASYSTTTTNNMGIMNFTTSGSGSTNSOGQTPQR 358
QY 307 -SGVLKDVVANIIEBOAKAAGEAKQAQIENNAQOKKYDEOQAKROBELKVSAGAGYL 365
DB 359 VSGLOGSAL-NIQONQISG-----SLQAGQKEGQNOQOQOQIILQP----- 403
QY 366 SGALILGGGIGVAVTAALHRKNQVPEQTTTTTTTTTTSARTVENKPAANTPAQGNVDP 425
DB 404 -QLVQGGQALQALQAA-----PLSGQFTTQALISOFTLQNLQALAVNSPIII- 451
QY 426 GSEDMESRSMATSTSTFPDTSIGT--VONPVADVETSLHDDQV--PFSNNTSQQ 480
DB 452 -----RTPTVGVNGOVSWQTLQNLQONPQACITTLAPMOCVSLGQTSSTNTLT 503
QY 481 NMGNTDSVVYSTIQHPPRDTTNGCARLLGNPSAGIQS--TYARLALSGLRHDMGGLTGS 539
DB 504 PLASAASTPAGTV-----TVNAQLSMP--GIQTINLSLGTSGIOVHIQGL----- 550
QY 540 NSAVNTSNPPAPGSH 555
DB 551 --PLAIAN--APGDH 561

RESULT 10

US-10-471-571A-2278
Sequence 2278, Application US/10471571A
Publication No. US20060115490A1
GENERAL INFORMATION:
APPLICANT: CHIRON SPA
TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS
FILE REFERENCE: P026927MO
CURRENT APPLICATION NUMBER: US/10/471,571A
CURRENT FILING DATE: 2003-09-12
PRIOR APPLICATION NUMBER: GB-0107661.1
PRIOR FILING DATE: 2001-03-27
NUMBER OF SEQ ID NOS: 5642
SOFTWARE: SeqWin99, version 1.03
SEQ ID NO 2278
LENGTH: 2478
TYPE: PRF
ORGANISM: Staphylococcus aureus
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (1)..(2478)
OTHER INFORMATION: FmcB protein [Staphylococcus
US-10-471-571A-2278

Query Match 5.2%; Score 149; DB 6; Length 2478;
Best Local Similarity 18.5%; Pred. No. 2.2; 265; Indels 198; Gaps 27;
Matches 128; Conservative 101; Mismatches 265; Indels 198; Gaps 27;

QY 9 NPNVNSIPPAAPLPSSQTD-GAGRGOLI-----NSTGLPSR 45
DB 1511 NTKAKNALQALDIPQASIDVKTNAABELTEMQKITELINNETTEKGDIGV--R 1568
QY 46 ALFTEPVNS-----MADSGDNRASDPGLPVNPMRLAAS-----EI 81
DB 1569 AAYEEGLNNINATTTGDTTAKDTAVGQVQQLHNPVVKPAGKELDQAAADKKTQIBQ 1628
QY 82 TLNDGEVLHD-HGPDLTNRQIGSSVFRVETQ-----DGKHAVGQRNGEVSVYL 133
DB 1629 TPNASQOEINDAKQEVDELNQAKTNVDSSTNEVYDNAVKKGK--AKINAVKTF-- 1681
QY 134 SPOEYARLOSIDPEGDKVFVTCGRGAGHAMVTVAADITEARQRIELELE----- 184
DB 1682 --SEYKK--DALAKIEDAVYAKVNEADNSASTSSEIAEKQKLAELKQTAQONVQA 1735
QY 185 -----PKGTGSKGAGESKVGELRESNSGAENTTETQTSST 222
DB 1736 TSKDDIEVQIHNLDNINDYTIPTGKKS--ATTDLVAVADQKNNISAD-TNATQDEKQ 1793
QY 223 SSLRSDPKMLALGTVATGL-----IGLAATGIVQALALTPPEPDSPTTDPAAA 272
DB 1794 AIKQVDQNVOTLAEISINNGVDGVDADALTGKKAIDAIQVDA-TVKPVANOAIIEVKAD 1852
QY 273 SATETATBDQLTKEAFQNPDNQKVNIDELGNAIPSGVLKDVVANIIEBOAKAGEBAKQ 332
DB 1853 TRESIDQSDQLTAE--EKTEALAMIKQITTDQAKGITDATTAAV-EAKAQGLEBAFN 1908
QY 333 AIENNAQOKKYDEOQAKROBELKVSAGAGYLSGALILGGGIGVAVTAALHRKN----- 387
DB 1909 IQIDSTERQKAIIELEFALDQ-----IEGVVNNADATTEKEAFNVA 1951
QY 388 -QVPEQTTTTTTTTTTTSGAR--TYENKPAANTPAQGNVDPSEDMESRR----- 435
DB 1952 LEDIISKATEDISDQTTAAELATVNSALBQLEKQ-RINPEVKNALEAIREVNVKQIEI 2010
QY 436 --SSWASTSTFFDTSIGTQONPVADV--KTSLL-----DSQVPTSNSNT 477
DB 2011 IKNADADSAKEIATDQGRFRPADKLDKIQTAFAELQNTTTPAIEALVPPQNDPA 2070
QY 478 SVQNGG--NTDSVVYSTIQHPPRDTTNGCARLLGNPSAGIOSTYARLALSGLRHDMGGL 535
DB 2071 NDTNNGIDNNDATANSNANATPENT-----GQPNVSETTANGKADASPTTPNNSDA 2122
QY 536 TG-----GNSAVNTSNPP 550

Db 2123 TGETTATSATDANDKPOANNSSVDASTNSP 2154

RESULT 11
US-11-342-171-6
Sequence 6, Application US/11342171
Publication No. US20060111558A1
GENERAL INFORMATION:
APPLICANT: Wild, Martha A.
APPLICANT: Cochran, Mark D.
TITLE OF INVENTION: RECOMBINANT INFECTIOUS LARYNGOTRACHEITIS VIRUS
TITLE OF INVENTION: AND USES THEREOF
NUMBER OF SEQUENCES: 72
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/11/342,171
FILING DATE: 27-JUN-2006
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/468,190
FILING DATE: 06-JUN-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/126,597
FILING DATE: 24-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 39116-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 985 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-11-342-171-6

Query Match 5.2%; Score 147; DB 7; Length 985;
Best Local Similarity 20.6%; Pred. No. 0.92;
Matches 121; Conservative 78; Mismatches 264; Indels 124; Gaps 22;

QY 20 PPLPSQDTGAGRGGLINSTGSLRALTFTPRN-SMADSDP-----NRASDVP-GLPV 71
DB 242 PVLKRVSSGF-----LVKYPIDIDRANMIVNIVANSSPADSGSVLAFTAREBGLSAIQ 296
QY 72 NPMRLAASEITLNDSEVLHD-HGPLDLINRQIGSSVFVEVTOEDGKH1----AVGQRNG 126
DB 297 HRIDMSGR---PRTETTFDQCKMIFETRYALGNSVPRDSDIRGATLPPPTDAAPPD 353
QY 127 VETSVLSDQEVARLOSIDPEGKDFVFTGSGGAGHNAVYASDITTEARQILLELPK 186
DB 354 TGTSTPTTVPEPAITTLIPRSTSDMGFSTARATGSETLSVPVQETDRLSTPLTLPL 413
QY 187 GTGSEKGA-GESEKGVGELRESNGAENTETOT-----STSTSLRS-DPKLWLA 234
DB 414 TPGESENLPFTTAPGISTETPSAAHETTQTQSAETVFTQSPSTESETARSQSOSEPMWY 473
QY 235 LGTVATGLIGLAATGIVQNAL---TPEDSPPTTDDPA----- 271

Db 474 TOTSTEOALALTOIATEALFTQTPSAEQMTFTQTPGAEAPQTPSTIPEITQSR 533
QY 272 ASATEIATRODTUEAFQNPONOKVINDELGNALPSCVLKDDVYANIE----EQAKAGE 327
DB 534 STPEETAPAPSADEVFTQSS---TVTEVFTQTPSTVPKTTLSSSTPEALFTTQSAQT 590
QY 328 EAKQOAIENNAQAOKKYDEQAKRQELKVSSGAGYLSGALLIGGIGVAVTAALHRKN 387
DB 591 EA-----FQTSASAPDPMRQSTETHPTQA-----PSTVPKATQTPS 629
QY 388 QPVEQTITTTTTTTTSARIVENKP-ANMTP-AQGNVDPGESEDTMESRRSS-MASTIST 444
DB 630 TEPEVLTPSTEPVPEFTRTIGAPEITQTPSAAPVYTRSSSTMPETAQSTPLASQNP 689
QY 445 FFDTSIGTVONPYADVTKSHDSQVTPNSNTS----- 478
DB 690 SSGTGTNTEPRTY-PVQTPHTQKLYTENKTLSPFVIVSEFHEMSTAESQTPLLDVKI 748
QY 479 ---VQNMGTDSVYVSTIQHPPR-----DTTNGARLLGNPSAGI 515
DB 749 EVKESNDGEVATGCVSTYKSPYRVETMKKVDLVDMDEISGNSPAGV 795

RESULT 12
US-11-342-171-66
Sequence 66, Application US/11342171
Publication No. US20060111558A1
GENERAL INFORMATION:
APPLICANT: Wild, Martha A.
APPLICANT: Cochran, Mark D.
TITLE OF INVENTION: RECOMBINANT INFECTIOUS LARYNGOTRACHEITIS VIRUS
TITLE OF INVENTION: AND USES THEREOF
NUMBER OF SEQUENCES: 72
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/11/342,171
FILING DATE: 27-JUN-2006
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/468,190
FILING DATE: 06-JUN-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/126,597
FILING DATE: 24-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 39116-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
INFORMATION FOR SEQ ID NO: 66:
SEQUENCE CHARACTERISTICS:
LENGTH: 985 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-11-342-171-66

Query Match 5.2%; Score 147; DB 7; Length 985;

Best Local Similarity 20.6%; Pred. No. 0.92;
Matches 121; Conservative 78; Mismatches 264; Indels 124; Gaps 22;

QY 20 PPLPSQTDGAGGQGLINSTGPLGRALFTPVNR-SKADSGD-----NPAADVP-GLPV 71
DB 242 PVLNRVSDGF-----LVKYPDPIDGRAMIVIANYSPADSGSVLAFTAFREGLPSAIQL 296
QY 72 NPMRLASEITLNDGEFVLHD-HGPLDTLNRIOGSSVFRVETOEDGKH-----AVQGRG 126
DB 297 HRIDMGTE---PFGTEITTDCCQMIETFPYRAGSNVPRDSDIRPGATLPPFPDPAAPDF 353
QY 127 VETSVVLSDDGEVARLOSIDPEGRKDFVTCGRGAGHAMVTVASDITEARQRIILELPEK 186
DB 354 TGRSPPTVPEPAITLIPRSTSDMGFPSTARATSEILSVVQETDRLLSTPLPL 413
QY 187 GCGESGGA-GESEGVLEBSNGAENTETOT-----STSTSLRS-DPKWLTA 234
DB 414 TPGESENTLFPPTAPGISTETPSAHHTTQSGAEVTFQSPSTESSTARSGQEPWYF 473
QY 235 LGTVATGLIGLATGIVQALAL---TPEDSPITTDDBDA----- 271
DB 474 TOTPSIOALTOQIAETALFTQPSAEQMTFTQPGAEPAQTPSTIPEFTQSR 533
QY 272 ASATETATRDQLTKEAFQHPDNQKVIDELGNALPGVLRKDVANIE---EQAKAGE 327
DB 534 STPEPTARAPSAPEVFTQSS---TVTEVFTQTPSTVPEKTLSSSTEPALFTRTOSAGT 590
QY 328 EAKQQAENNAQAKKYDEQQAQROBELKVSGAGYGLSGALLGGIGVATVTAALHRN 387
DB 591 EA-----FTQSSAEPPTMRQSTETETFTQA-----PSTVKATQTPS 629
QY 388 QPVEQTTTTTTTTTSARTVENKP-ANNTP-AQGNVDPGSDTMESSRS--MASTST 444
DB 630 TPEPVLTQSPSTPEVFTPLGAEPEITQTPSAPEVTRSSSTMEETQSTPLASQNP 689
QY 445 PFDTSIGTVQNPYADVKTSLHDSQVPTSNSNTS----- 478
DB 690 SSGTGHNTPEPRY-PVQTPHQKXYTENKTLSPFTVVSERHEMSTASQTPLLDVKI 748
QY 479 ---VQMGNTDVSYSTIQHPPR-----DTTNGCARLLGNPSAGI 515
DB 749 EVKFSNDGEVTATCVSTKSPYKVEETNMKVDLVDVDEISGNPAGV 795

RESULT 13
US-11-013-711-9
; Sequence 9, Application US/11013711
; Publication No. US20060159709A1
; GENERAL INFORMATION:

; APPLICANT: Loomore, Sheena M.
; APPLICANT: Sasaki, Ken
; APPLICANT: Yang, Van Ping
; APPLICANT: Klein, Michael H.
; TITLE OF INVENTION: RECOMBINANT HIGH MOLECULAR WEIGHT MAJOR OUTER MEMBRANE
; FILE REFERENCE: 1038-921MIS-jb
; CURRENT APPLICATION NUMBER: US/11/013, 711
; CURRENT FILING DATE: 2004-12-16
; PRIOR APPLICATION NUMBER: US/09/361, 619
; PRIOR FILING DATE: 1999-07-27
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 2053
; TYPE: PRT
; ORGANISM: Moraxella catarrhalis
US-11-013-711-9

Query Match 5.1%; Score 145.5; DB 7; Length 2053;
Best Local Similarity 19.9%; Pred. No. 2.8;
Matches 110; Conservative 82; Mismatches 218; Indels 143; Gaps 22;
QY 116 GGHIAVG-----QRNGVETSV-----VLSDDGEVARLOSIDPEGRKDFVFTGGRG--- 159

DB 273 GREGVAGFSGQILDRNNNDASAYVPLGKTLAQYRATRQ-----GDSITDFSIGSNNN 327
QY 160 -----GAGHAMVTVASDITEARQRIILELPE-----KQGESKAGSGKV 200
DB 328 NSSIRRKITVAG-----SRDPAVNVAAQLKVEELANRKITFKQDGDNNNSVERGL 381

QY 201 GEIARENSGAENTETOTSTSTSLRSDPRLMALGTVATGLIGLATGIV----- 251
DB 382 GNTLTTKGDAQ--TNALTEANIGVTVDNGKLKAKELGLTSVSAITKITYSNTNNN 439
QY 252 ---QALATPEPDSPTTDP-----DAASATEATARDQLTKEAFQNP----- 291
DB 440 AELQSGGLFSPITGKTIDKTYVSIIDGLKFTNDSNIAIKTTRITKKKIGFAGTNDGVD 499
QY 292 -----DNQKVIDELGNA-IPSGVLK-DDVANIEEQAKAGEAKQAENNA----- 337
DB 500 ESKPYLDNEKLKV---GNSTLNSGSLTVANNNGNKIQVGANG--IKPATVANNVANTS 554
QY 338 -----AQAQKKYDEQQAQROBELKVSGAGYGLSGALLGGIGIV 377
DB 555 TVGTARITEKIGFAGTNDGVDQAPYLDKERLKVGRVEITTDGGINAGNHKITGLTNGI 614
QY 378 AVTAALHRKQNPVEQTTTTTTTTTSARTVENKP-ANNTPAQGNVDPGSDTMESSRS 437
DB 615 ANTDAV-----TIKQKAKPFLNAGDGISSINNGDLVDSSGNTTTPYNTISVKTJKLN 669
QY 438 MASTS-----STFEDTSIGTVQNP--PYADVKTSLHDSQVPTSNSNTSVQMGNTDVS 488
DB 670 SNGTSGNNKFSYNMADNNSLVTAQDLADYLNKVNETAQALP-----SPRYQGDNSNA 725
QY 489 VSTIQHPPRDTTNGCARLLGNPSAGIQSTVARELALSGLRHDMG---GLTGSNSAVN 544
DB 726 I--TVGKQDNGKTFPTLKLKGEVGVITNRAATGVTFPIDQSNGLTTPKLTGVSPTNGN 783
QY 545 ---TSNNPAPAS 554
DB 784 RLIVIEQVPSADGN 796

RESULT 14
US-11-013-711-11

; Sequence 11, Application US/11013711
; Publication No. US20060159709A1
; GENERAL INFORMATION:
; APPLICANT: Loomore, Sheena M.
; APPLICANT: Sasaki, Ken
; APPLICANT: Yang, Van Ping
; APPLICANT: Klein, Michael H.
; TITLE OF INVENTION: RECOMBINANT HIGH MOLECULAR WEIGHT MAJOR OUTER MEMBRANE
; FILE REFERENCE: 1038-921MIS-jb
; CURRENT APPLICATION NUMBER: US/11/013, 711
; CURRENT FILING DATE: 2004-12-16
; PRIOR APPLICATION NUMBER: US/09/361, 619
; PRIOR FILING DATE: 1999-07-27
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 2314
; TYPE: PRT
; ORGANISM: Moraxella catarrhalis
US-11-013-711-11

Query Match 5.1%; Score 145; DB 7; Length 2314;
Best Local Similarity 20.8%; Pred. No. 3.5; Mismatches 248; Indels 138; Gaps 27;
Matches 125; Conservative 89;

QY 26 TDGAGRGQGLINSTGPLGRALFTPVNRSM-----ADSGDNRASDVPGIPLV 71
DB 245 TIAISNQAIN-----YGALLAGADTRVDLDVIGIALGVGSGQLNNNNNNNNNAYVY--EG 298
QY 72 NPMRLASEITLNDGEFVLHDHGPLDTLNRIOGSSVFRVETOEDGKHIAVGQRNGVETSV 131

```

Db      299  NGSNIKSKATNGNGF-----SIGSSSTIK-----RKININVA--GYEDTD 336
Qy      132  VLSDOEYARLOSIDPEGKDFVETGGRGAGHAMTVASDITEARQRLLELEPKGTCS 191
Db      337  AVN---VAQLKAVENIARQITFKGDNDGTG-----YKKKGELTITIKG--GET 380
Qy      192  KG--AGESEKGVGELNESNG-----AENTTETQSTSTSLRSDPKMLALGTAVATGIG 244
Db      381  QADKLTDDNNIGVVDNNITGLKVKLAKNLSGLT--VSTKNLTASEKV---TVSG-- 431
Qy      245  LAATGIVQALALTPEDSPPTTD-----PDAAASATETATRDQLTEAFQNP-- 291
Db      432  -NNVAELQSGGLTPTTNASTDKTVYGTDLKFTDNSNTALIEDTTRITKXIGFSNAG 490
Qy      292  --DNQKNID---ELGNAI--PSGVLKDDVVANIEEQAKAAGEAKQQAENNAQAQK 343
Db      491  TVDENKKPYLDKDKLVGNGSTLNNGGLTVNNITGGSNKOIOVGADGICFADVVVNSNAK 550
Qy      344  YDEQQAARQF-----ELKVSAGAGYGLSGALLIGGIGVAVTALHKNOP---VEGTT 394
Db      551  FGTTRITEEELGFADGKVDKKSPP--LDKKOLQVGVKAITKDSGINAGDQKISNVKAT 609
Qy      395  TTTTTSARTVENKPPANNTPAQANVDTPGSEDTMESRRSSMASTSTPFDTSIG-- 452
Db      610  DDTDAVTKQLKQVQODADGALQSFIRDEKQGEFTI--SNLYSNGNTPTTFETIFAGEN 668
Qy      453  --TYVNPADYKTSIHLDSQV---PTSNSNTSVQNNGNTDSVYVSTIQHPRPRTTNGARL 507
Db      669  GISISNDIAKGVKVGIDPINGLETTPKLTVGSDKDGKTQVLVE---QVASGNDTYNIIR- 724
Qy      508  LGNPSAGIOSTFYARLALSGLRHDMGGLTGCS-----NSAVNTSNPPAPG 553
Db      725  -----GLSPITPSTINAGCVTTTEQGNITTSDEKSKRAASIGDILNTGPNLKANSNVG 778

RESULT 15
US-11-056-355B-87568
; Sequence 87568, Application US/11056355B
; Publication No. US20060150283A1
; GENERAL INFORMATION:
; APPLICANT: Brover, Vyacheslav
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
; FILE REFERENCE: 2750-1590FUS2
; CURRENT APPLICATION NUMBER: US/11/056.355B
; PRIOR FILING DATE: 2005-02-14
; PRIOR APPLICATION NUMBER: 60/544,190
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 119966
; SEQ ID NO 87568
; LENGTH: 804
; TYPE: prt
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
; LOCATION: (1)..(804)
; OTHER INFORMATION: Ceres Seq. ID no. 12701803
US-11-056-355B-87568

```

Query Match 5.0%; Score 143; DB 7; Length 804;

Best Local Similarity 18.7%; Pred. No. 1.2; Indels 218; Gaps 26;

Matches 121; Conservative 87; Mismatches 221; Indels 218; Gaps 26;

```

Qy      17  PPAPLPSCOTDAGRGQILNSTPLGSRALFTPVNNSMADGDNRASDPGLPVNPMRL 76
Db      12  PPPPPPPSRT-----VVVALSGSSKSKYVVTWAIKKFATEG----- 47
Qy      77  AASEITLNDGFEVLADHGLDLTLNROIGSVFRVETQED-----GK 117
Db      48  -----NVGFLLHHPWITSVPTPMGNALPISSEVRDVTVAQROELIMOSEMLKPYT 100

```

```

Qy      118  HIAGORNGVETSVLSPDOEYARL--QSIDPEGKDFVETGGRGAGHAMTVASDITEAR 176
Db      101  KLPFRKRAVAVVYIESNNVAAALAEVTRSIDRIYV-----GSSSRFFFRKADICS-- 154
Qy      177  QRILELEP-----KGTSEKGAAGESKGVLELR-----SNSAENTTETQST 220
Db      155  --VISALMPNFCVYVVSKGTLSCVRPESDSDGNATTIREDGSERTNSSSGSPISGQASV 212
Qy      221  --STSLRSDPKMLALQTVATGLIGLAATGIVQALALTPEDSPPTTDPPAA--ASATET 277
Db      213  PMETSVSGSDTRCQSL-----DAEARVDSINRSSTDY 247
Qy      278  ATR-----DQLTKEAFQNPDNQKVNIDELGNAI--SGVLKDDVVANIEEQAKAAG 326
Db      248  TSWTTPRRRDVEERKEAMSSSSNR-----EYGNIGTRSWSGMGVDITHSRAAQASNMS 303
Qy      327  EBAKQQAENNA-----QAQKQYD-----QAQKR 351
Db      304  DALSEQSYTDNQVNLNFEVEKLRALRHVQEMVAVAQTEFPDASRKLQELNQRLLEAIK 363
Qy      352  QEELKVSAGAGYGLSGALLIGGIGVAVTALHKNQVPEQTTTTTTTTTSARTV--- 408
Db      364  LEEILKLEYARLEA-----EKEKQNFEXARDAESMRERAELEIQR 406
Qy      409  ---ENKPPANNTP---PAQANVDTPGSEDTMESRRSSMASTSTPFDTSIGTYVNPADY 461
Db      407  REAEKKSARDTKEKKKLEGTIGSPLOLYQHFAWEIIMATSS--FSEELKIG--MGAYGAV 463
Qy      462  -KTSIHLDSQVPT---SNSNTSVQNNGNTDSVYVSTIQHPRPRTTDCARLILNPSAGIQS 517
Db      464  YKCNILHTTAVVVKVLOSAMQLSKQFOQEELEILSKIRHP----- 502
Qy      518  TYARLALSGLRHDMGGLT-----GGS--NSAVNTSNPPAPGSHRF 557
Db      503  ---HLVLLGACPEQGALVYEYENGSLIEDRLFOVNSPPLWFERF 546

```

Search completed: August 1, 2006, 22:29:38
Job time : 35 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using SW model

Run on: August 1, 2006, 22:08:25 ; Search time 42 Seconds

(without alignments)
1278.308 Million cell updates/sec

Title: US-09-189-415D-11

Perfect score: 2840

Sequence: 1 MFIGNLGHNPNVNNISIPAP.....SNSAVNTSNNPPAPGSHRFV 558

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2840	100.0	558	2 A98199	translocated intim
2	2840	100.0	558	2 E86045	probable transloca
3	200	7.0	2232	2 T34434	hypothetical prote
4	195.5	6.9	1229	2 T25697	hypothetical prote
5	187	6.6	1291	2 T13389	hypothetical prote
6	172.5	6.1	1630	2 A53577	ascities sialoglyco
7	172.5	6.1	2271	2 F90073	hypothetical prote
8	167.5	5.9	1192	2 T18611	probable serine/th
9	167	5.9	2468	2 A83412	hypothetical prote
10	166	5.8	796	2 T21460	hypothetical prote
11	164	5.8	971	2 B90835	hypothetical prote
12	164	5.8	973	2 C85693	probable tail fibe
13	164	5.8	1275	2 T33369	hypothetical prote
14	164	5.8	1770	2 A71517	hypothetical prote
15	163.5	5.8	1802	2 G89287	protein H39E23.1 l
16	163.5	5.8	1802	2 S69703	HKR1 protein precu
17	163	5.7	786	2 T16509	hypothetical prote
18	160	5.6	1122	2 G64887	probable tail fibe
19	160	5.6	1829	2 T24583	hypothetical prote
20	158.5	5.6	1063	2 D86731	hypothetical prote
21	157.5	5.5	3013	2 A80480	probable invasin y
22	156.5	5.5	3570	2 T45025	mucin MUC5B, trach
23	155.5	5.5	3507	2 T34513	hypothetical prote
24	152.5	5.4	1077	2 A44067	serine-rich protei
25	152	5.4	461	2 UN0097	secreted 45k prote
26	152	5.4	997	2 T43523	cutl7 protein - fi
27	151.5	5.3	2660	2 E85822	probable invasin z
28	151	5.3	918	2 T02759	hypothetical prote
29	151	5.3	1306	2 S25370	M82 protein - yea

30	150	5.3	888	2 T46726	secreted acid phos
31	150	5.3	1026	1 A40315	maternal effect pr
32	150	5.3	1034	2 JC2143	ice nucleation act
33	150	5.3	1460	2 D81675	polymorphic membra
34	149.5	5.3	1772	2 A45532	major merizole su
35	149.5	5.3	2514	2 F81045	hemagglutinin/hemo
36	148.5	5.2	1189	2 T51024	related to C2H2 zi
37	148.5	5.2	1892	2 S56852	hypothetical prote
38	148	5.2	2035	2 A40718	host cell factor C
39	148	5.2	3190	2 T13828	CREB-binding prote
40	147.5	5.2	797	1 VGBEX1	glycoprotein X pre
41	147.5	5.2	1547	2 T28657	blackjack protein,
42	147.5	5.2	2055	2 T31110	extracellular matr
43	147.5	5.2	2059	2 D82671	surface protein XF
44	147.5	5.2	3624	2 AD0835	large repetitive p
45	147.5	5.2	4558	2 C82199	RTX toxin RtxA VCI

ALIGNMENTS

RESULT 1

A98199
translocated intimin receptor Tir [imported] - Escherichia coli (strain O157:H7, substr:
C/Species: Escherichia coli
C/Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #ext_change 09-Jul-2004
C/Accession: A98199
R/Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 9, 11-22, 2001
A/Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen
A/Reference number: A99629; PMID:21156231; PMID:11258796
A/Accession: A98199
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-558 <HAY>
A/Cross-references: UNIPROT:Q9R396; UNIPARC:UPI00000D00CA; GB:BA000007; PIDN:BA837984.1,
A/Experimental source: strain O157:H7, substrain RIMD 0509952
C/Genetics:
A/Genes: EC84561

Query Match	100.0%	Score 2840;	DB 2;	Length 558;
Best Local Similarity	100.0%	Pred. No. 6.7e-150;		
Matches	558;	Conservative	0;	Mismatches 0;
			Indels	Gaps
QY	1	MFIGNLGHNPNVNNISIPAPPLPSQTDGAGRGQIINSTGPLSRALFTPVNSMADSGD	60	
DB	1	MFIGNLGHNPNVNNISIPAPPLPSQTDGAGRGQIINSTGPLSRALFTPVNSMADSGD	60	
QY	61	NRASDVPGIPVNPMLAASEITLNDGFEVLHDHGPLDLNROISSVFRVETQEDGKHIA	120	
DB	61	NRASDVPGIPVNPMLAASEITLNDGFEVLHDHGPLDLNROISSVFRVETQEDGKHIA	120	
QY	121	VGQRNGVTSVVLSPDOEVARLOSIDPEGKDKVFIFGGRGAGAHAVTASDITTEARQRL	180	
DB	121	VGQRNGVTSVVLSPDOEVARLOSIDPEGKDKVFIFGGRGAGAHAVTASDITTEARQRL	180	
QY	181	ELLLEPKGTGESKAGESKGVGLERBSNGAENTTETOTSTSSIRSPDKMLAAGTVAT	240	
DB	181	ELLLEPKGTGESKAGESKGVGLERBSNGAENTTETOTSTSSIRSPDKMLAAGTVAT	240	
QY	241	GLIGLAATGIVQALALTPEDPSPTTDDPAASATETATRDQLTEAFQPNPNQKYNIDE	300	
DB	241	GLIGLAATGIVQALALTPEDPSPTTDDPAASATETATRDQLTEAFQPNPNQKYNIDE	300	
QY	301	LGNAIPSGVLKDDVVANIEBQAKAAGEAKQOALENNQAQKRYEQQAQKEELKVSQG	360	
DB	301	LGNAIPSGVLKDDVVANIEBQAKAAGEAKQOALENNQAQKRYEQQAQKEELKVSQG	360	
QY	361	AGYGSLGALLIGGIGIVAVTAAALHRKNQPVETTTTTTTTTTSARTEYENKPAANTPAQG	420	
DB	361	AGYGSLGALLIGGIGIVAVTAAALHRKNQPVETTTTTTTTTTSARTEYENKPAANTPAQG	420	

Qy	421	NVDTPGSEDDTMEKRRSSMASTSTFPFDTSIGTYQNPADYAKTSLHDSQVPTSNMSTSVQ	480
	422		
	423		
	424		
	425		
	426		
	427		
	428		
	429		
	430		
	431		
	432		
	433		
	434		
	435		
	436		
	437		
	438		
	439		
	440		
	441		
	442		
	443		
	444		
	445		
	446		
	447		
	448		
	449		
	450		
	451		
	452		
	453		
	454		
	455		
	456		
	457		
	458		
	459		
	460		
	461		
	462		
	463		
	464		
	465		
	466		
	467		
	468		
	469		
	470		
	471		
	472		
	473		
	474		
	475		
	476		
	477		
	478		
	479		
	480		
	481		
	482		
	483		
	484		
	485		
	486		
	487		
	488		
	489		
	490		
	491		
	492		
	493		
	494		
	495		
	496		
	497		
	498		
	499		
	500		
	501		
	502		
	503		
	504		
	505		
	506		
	507		
	508		
	509		
	510		
	511		
	512		
	513		
	514		
	515		
	516		
	517		
	518		
	519		
	520		
	521		
	522		
	523		
	524		
	525		
	526		
	527		
	528		
	529		
	530		
	531		
	532		
	533		
	534		
	535		
	536		
	537		
	538		
	539		
	540		
	541		
	542		
	543		
	544		
	545		
	546		
	547		
	548		
	549		
	550		
	551		
	552		
	553		
	554		
	555		
	556		
	557		
	558		
	559		
	560		

```

RESULT 2
E86045
probable translocated intimin receptor protein tir [imported] - Escherichia coli (strain)
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C:Accession: E86045
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew,
Miller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Diallanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: E86045
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-558 <STO>
A:Cross-references: UNIPROT:O9R396; UNIPARC:UPI00000D00CA; GB:AE005174; NID:g12518449; E
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: tir

```

Query Match	100.0%;	Score 2840;	DB 2;	Length 558;
Best Local Similarity	100.0%;	Pred. No. 6.7e-150;		
Matches 558;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0

QY	1	MPIGLGNHPVNNNS	IPAPAPLP	PSOTDAGRGQ	INSTG	PLGSAL	FTPRNSMA	SGD	60	
Db	1	MPIGLGNHPVNNNS	IPAPAPLP	PSQTDAGRGQ	INSTG	PLGSAL	FTPRNSMA	SGD	60	
QY	61	NRASVPGLPVNPM	LAASEITL	NDGFEVL	HDHGL	DLTLNR	QIGSSV	RVETOE	DGKIIA 120	
Db	61	NRASVPGLPVNPM	LAASEITL	NDGFEVL	HDHGL	DLTLNR	QIGSSV	RVETOE	DGKIIA 120	
QY	121	VGORNGVETS	VVLSDOE	YARLOS	IDPEG	KDFEF	FGRG	GAGHAW	TVASDITTEARQIL 180	
Db	121	VGORNGVETS	VVLSDOE	YARLOS	IDPEG	KDKVF	FGRG	GAGHAW	TVASDITTEARQIL 180	
QY	181	ELIEPKGTG	ESKGAGES	KGVGEL	RESNSGA	ENTTETOT	STSSLR	SPDKL	MLALGIYAT 240	
Db	181	ELIEPKGTG	ESKGAGES	KGVGEL	RESNSGA	ENTTETOT	STSSLR	SPDKL	MLALGIYAT 240	
QY	241	GLIGLAAG	IGIYQAL	LTPEPDS	PTTTD	PPAASA	FETATRO	LTTEA	QONPNQKVNIDE 300	
Db	241	GLIGLAAG	IGIYQAL	LTPEPDS	PTTTD	PPAASA	FETATRO	LTTEA	QONPNQKVNIDE 300	
QY	301	LGNALPSG	VLKDDV	ANIEE	OKAAGEA	EAKQOAI	ENNNQA	OKKYEO	QAKQOEELKXSSG 360	
Db	301	LGNALPSG	VLKDDV	ANIEE	OKAAGEA	EAKQOAI	ENNNQA	OKKYEO	QAKQOEELKXSSG 360	
QY	361	AGYGLS	GALLIG	GIGIV	AVTAA	LHKKNO	PVEQ	TTTTTTTT	TSAR	YENKPNANTPAQG 420
Db	361	AGYGLS	GALLIG	GIGIV	AVTAA	LHKKNO	PVEQ	TTTTTTTT	TSAR	YENKPNANTPAQG 420
QY	421	NVDTPG	SEDTMES	RSSMA	STSTPE	DTSSIG	TVON	PYADV	TS	LHDSQVPTSNSNTSVQ 480
Db	421	NVDTPG	SEDTMES	RSSMA	STSTPE	DTSSIG	TVON	PYADV	TS	LHDSQVPTSNSNTSVQ 480
QY	481	NMGNTDS	VVYYS	TIHQ	PRD	TTNGAR	LGN	PSAG	QSTY	ARALASGLRHDMGGLTGSN 540
Db	481	NMGNTDS	VVYYS	TIHQ	PRD	TTNGAR	LGN	PSAG	QSTY	ARALASGLRHDMGGLTGSN 540

QY	541 SAVNTSNPPAPGSHRFV	558
Dd	541 SAVNTSNPPAPGSHRFV	558

RESULT 3
T34434
hypothetical protein K06A9.1a - *Caenorhabditis elegans*
C|Species: *Caenorhabditis elegans*
C|Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C|Accession: T34434
R|Geisels, C.; Gattung, S.
submitted to the EMBL Data Library, December 1996
A|Description: The sequence of *C. elegans* cosmid K06A9.
A|Reference number: Z21525

A1status: preliminary; translated from GB/EMBL/DBJ
A1Molecule type: DNA
A1Residues: 1-2232 <GI>
A1Cross-references: UNIPROT:Q81FX6; UNIPARC:UPI0000086219; EMBL:U08046; PIDD:AACT0890.1;
A1Experimental source: strain Bristol N2; clone K06A9
C1Genetics:
A1Gene: CESP:K06A9.1a
A1Map position: X
A1Introns: 38/1; 75/3; 103/3; 132/2; 158/2; 222/1; 1088/1; 1367/1; 2039/1; 2049/1; 2075/

Query Match	7.0%;	Score 200;	DB 2;	Length 2232;
Best Local Similarity	21.2%;	Pred. No. 0.0056;		
Matches 127;	Conservative 80;	Mismatches 278;	Indels 114;	Gaps 21;

QY 9 NPNVNNSIPPAPPLPSQTDGAGGRGQLINSTGPIGSRALFT'PVRNSM----- 55

Db 1343 SPSSISPVPPTSSPIPTTFASSTSGSTISDVSSVSTSL-APLSSSLPSTVPSSSTQSFSS 1401

QY 56 ADGDNRASDV-----GLPVNPMRLASEITLNDGFVLDHGPDLTLNRQIGSSVER 109

```

Db      1402 TSESSKASSSPVPSQTSTPTNPTGSTEESSTLLSSTISGTOH--TTMSKASSGSTSP 1458

```

QY 110 VETQEDGKHIAVG--QRNGVETSVLSDQFYARLQSIDPEGKDFVFTGGRGAGHA--- 164

Db 1459 STNSQTGSTVTMGSSSTSGVSTSSASSTQPQ MSTSQGSSAGSTVASSTASPAASSTAPSS 1518

QY 165 -----MTVASDITEARQRIELLEPKGTGSGGAGESKGVGELRESNSGAENTTET 216

Db 1519 TGTMSSTSSGCTVCGTISES-STTASASSQTGCTVTMGSSSTSGV-----STSSASSTOP 1571

QY 217 QTSTSTSLRSDPKLWLALGTVAATGLIGLAATGIVQALALTPEPDSPTTTDDPAASATE 276

Db 1572 QMSTQGS-----SAGSTVASSTAGLVSTV-----PSSTGTMGSTSSGTVGS 1615

QY 277 TATRDÖLTKEAÖNPDPNÖKNIDELGNAIPSGVLKDDVVANIEEÖAKAKAGEEAKÖÖAIEN 336

```

Db      1616 TISESSTASA---SSQTGSTVTMGSSSTSGV-----STSSASSTQPQMTSQ 1660

```

337 NAQAQKYDEQAQRQEELKVSSGAG-YGLSGALILGGIGVAVTA--ALHRKNQPEQT 393

Db 1661 GSSAGSTVASTTGLVSTSTVPSSTGTMGSTSSGTVGSTISESSTAASASSQTGSTVTMG 1720

QY 394 TTTTTTTTSA-----RIVENKPAANTPAQGNVDTPGSEDTMESRRSSMASTSTFF 446

```

Db      1721  ::::: 1776
          SSSTGCVSTSSASSGQPQMSTSGGSSAGSTVVSSTA-SPAASSTAPSSGTMSSTSS---

```

QY 447 DTSSIGTV-----QNPYADVKTSLHDSQVPTSNSNTSVQNMGN-----DSVVYSTIQH 495

Db 1777 -----GTGGTMSQSTASTTSHTGSTVTLGSSSTSNQMSTSQGSSVGSTVASSTAGL 1831

QY 496 PPRDTTNGARLLGNPSAGI-----QSTYARLALSGGLRHD MGGLTGGSNSAVNTSN 547

Db 1832 VSTSTVPSTGTMGSTSSGTVGSTISESSTTASASSQTGTVMG---SSSTSGVSTSS 1887

AKB0001 3
T25697

hypothetical protein F16F9.2 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: J25697
R:Fulton, B.
submitted to the EMBL Data Library, August 1996
A:Description: The sequence of C. elegans coemid F16F9.
A:Reference number: Z20071
A:Accession: J25697
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1229 <FULL>
A:Cross-references: UNIPROT: O94185; UNIPARC:UPI00001786FF; EMBL: U67556; PIDD: AAB07691.1.
A:Experimental source: strain Bristol N2; clone F16F9
C:Genetic8:
A:Gene: CESP.F16F9.2
A:Map position: X
A:Introns: 35/1; 361/1; 384/1; 482/2; 517/1; 971/1; 1021/2; 1179/3

Query Match	6.9%	Score 195.5	DB 2	Length 1229
Best Local Similarity	22.0%	Pred. 0.0044		
Matches	90	Mismatches 155	Indels 111	Gaps 13
QY	167	TVASDITEARQRIELLEPKGTGSEKSGAGSEKGVGELR-----ESNSGAENTLETOTSTYT	222	
DB	597	SVSTEKSTYTKKASTTEETPTTDEPTTTTESSTTGKATPELSTSEETTTTELKLTTEG	656	
QY	223	SSLASDPRLMLALGTVAITGLIGLATGIVQALALTPEPDSPTTTDPPAASATETATRDQ	282	
DB	657	STTTEEP-----TTTAIFPAESTGII-----TTDEBTSTSTTPEITSTKEIVESA	704	
QY	283	LTKFAF-----QNPDNQKVAIDDELGNALPISGVLLADV-VANIEBOQAAGEAKQ	332	
DB	705	ITQTSVSVSESTPRQLPERKCALVNRKKNHLE-VLKEKRRLKXKSTSTTSDSS--	760	
QY	333	AIENNAQAQKRYDEQQAQROEKLKVSAGYIGLSGALILGGIGVATPALHRKNQVPEQ	392	
DB	761	--ETTVAVENIDVTTTTEKEKV-----VQTPITTEKSTTOEE	797	
QY	393	TTTTTTTTTTSARTVENKPPANTPTPAQGNVDTPGSEDTMESGRSMATSTSPFDTSIG	452	
DB	798	TTTTTTTTTEKTSKTTTEKPTTSSA-----TTETTTSEPTTSETTYDTSAT	846	
QY	453	TVQNPPVA-DVKTSLHDSQVPTSNS-----NTSVQNMG-----	483	
DB	847	TEESTTAETTTTTSATSETTSSAALFITESPENTALQSSQKSEENESSAEKPGAR	906	
QY	484	-----NTDSVYVYITIQHPRNTDNGARLLGNPAGIOST	518	
DB	907	DPVPPKAKTIVKPAETTSAAVAASATTTTEPTTTEKSTTLETPP--LEAT	953	

RESULT 5
 T13389
 hypothetical protein 115C2.10 - fruit fly (*Drosophila melanogaster*)
 C:Species: *Drosophila melanogaster*
 C>Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 09-Jul-2004
 C:Accession: T13389
 R:Seiler, C.; Valenči, P.; Darlamićson, A.; Henderson, N.; Campbell, L.; Glover, D.
 Submitted to the EMBL Data Library, May 1999
 A:Description: Sequencing the distal X chromosome of *Drosophila melanogaster*.
 A:Reference number: 217665
 A:Accession: T13389
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1291 <CAT>
 A:Cross-references: UNIPROT:O77261; UNIPARC:UP10000078DB; EMBL:AL031581, NID:e1320978;
 C:Genetics:
 A:Cross-references: FlyBase:FBgn020381
 A:Map position: X
 A:Introns: 238/3, 1225/1
 A:Note: EG:115C2.10

Query Match	6.6%;	Score 187;	DB 2;	Length 1291;
Best Local Similarity	20.4%;	Pred. No. 0.014;		
Matches 130;	Conservative 74;	Mismatches 245;	Indels 188;	Gaps 25;

```

0Y 5 NIGHNPNVNNISIPAPPLPQSOTDQ---AGGEGOLINSTPGSRALPFPVNNSMADSGDN 61
Db 438 NSTNSNSNTMTDSTGCPSETSTINGLVASGGAG-----GATGAAMLPTP---SQOSTGK 488
0Y 62 PAS-----DVPGLPVNPMRLA-----ASEITLNDGEVLHDGPLEDTLNRQ 102
Db 489 EATAVSLLEKKLTPNVVVSPLTWKELRQKMTKYDAEIMTANAAVQOOHH----- 539
0Y 103 IGSSVERETQEDGKHIAVGGRNVEVSVLSDDEVALQSIDPBGMKQFPTGGRGAG 162
Db 540 --QHFFHHHHHHHHHHHNGHASTGAELATAVQOMAMQPRG-----VGTGAAG 588
0Y 163 HAMVTVASDITEARORILELLEPGGTGESKAGESKGVELRESGAEMLTETOTSTST 222
Db 589 NAGATVTVSSVA-----AGASEVNVGGRSTSLRKMVRNS-----TSESI 627
0Y 223 SSLSDPLKMLALGTVAATGLIGLATGCIYQALALTPEBDSFTTTDPPDAASATTATRDQ 282
Db 628 STASADEVT-----APVVAASISLPSKAPVVLMPCKPQAMAIALHQ 670
0Y 283 LTKSAFQPNQKYNIDELGNALPSGVLKPDVNVANIBEQAKAGEBKQO-----A 333
Db 671 SQOQLRSEKQCKLTD-----GSSSD---TSSEQCK---EQKODHQLPCKMFS 716
0Y 334 IENNAQOKKYDEQOAKROBELKVSAGAGVGLSGALLGGSIGVAVTAAHLRKNQPYEQT 393
Db 717 LAEPQPEKS-EEQOQEQCKRVNTNSAGRVGL-----VARIATAHNNN-----IA 760
0Y 394 TTTTTTTTTTSARV-----ENKPANPTAQGVNDTPGSEDMTESRRSSMASTSTFPDTSS 450
Db 761 TTTSSSSSNKATTITTCNNHNNSNNSRINHNSMLSRYSKRPAPSEASSIPSSSS 820
0Y 451 IGTVO-----NPAVADVKTSLHDSQVPTSNSNTSVOMMGVTDVSVVSTIOH----- 495
Db 821 ENQOQOATRRSCSTPAVKKNLLASFPDDPPTSGTIGIEQLKDESVYTSYPVKQKSRRAA 880
0Y 496 -----PRDITTDNGARL-LGNPSAGIOST-----YARLASGG 527
Db 881 LAAAQSIHCEALGFPPTGSGTCSQKRAQAGEPITCSSTIISNVEPLKTPBERLKTILR 940
0Y 528 LRH-----DMGGLTGSGNSANVTSSNNPPAPSHR 556
Db 941 MKRSPILDEVIELG-TSLNSGAGRG---APGSHR 971

```

RESULT 6
A53577
acities sialoglycoprotein 1 - rat (fragments)
C:Species: Rattus norvegicus (Norway rat)
C:Date: 12-Apr-1995 #sequence_revision 12-Apr-1995 #text_change 07-Feb-1997
C:Accession: A53577
R:Wu, K.; Fregien, N.; Carraway, K.L.
J. Biol. Chem. 269, 11950-11955, 1994
A:Title: Molecular cloning and sequencing of the mucin subunit of a heterodimeric, bifurcated
A:Reference number: A53577; MUID:94216302; PMID:8163496
A:Accession: A53577
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1630 <MUA>
A:Cross-references: UNIPARC:UPI000017C8BB; GB:U06746
C:Keywords: glycoprotein

Query Match 6.1%; Score 172.5; DB 2; Length 1630;
Best Local Similarity 20.0%; Pred. No. 0.12;
Matches 136; Conservative 70; Mismatches 272; Indels 203; Gaps 23;

OY 12 VNNSTPAP----PLPSQDTGAGGKQLINSTEPLSGRLFTYVRKSMADSGNRASDVP 67
db 28 VNNSTTAPKTSALPSTNPSTNPQMTSGVSNPTL-----SSRYRNTKNGQASPMWTSSIT 82

A:Gene: CESP:H39E23.1a; CESP:H39E23.1b
A:Map position: 5
A:Introns: 10/3; 33/3; 67/3; 139/2; 183/3; 451/3; 487/3; 535/3; 631/1; 825/2; 914/3; 992
C:Species: alternative splicing; APT: phosphotransferase; serine/threonine-specific pdc
F:1-1192/Product: probable serine/threonine-specific protein kinase, long splice form #8
F:1-487,336-1192/Product: probable serine/threonine-specific protein kinase, short splice
Query Match 5.9%; Score 167.5; DB 2; Length 1192;
Best Local Similarity 20.3%; Pred. No. 0.15;
Matches 130; Conservative 68; Mismatches 256; Indels 185; Gaps 23;
QY 36 INSTGSRALFTPVNSMADSGDNASDPGLPVNPKMLAASEITLNDGEVFLHDHGP 95
DB 505 INVSSSLGQHPAGVITREHYT-----SSASGSSASPSRYSRSATAT-GASITGSL 557
QY 96 LDTLNQ-----IGSVRVETOEDGKILANGQNGVTSVYLSDOEVARLOSID 145
DB 558 ASAAANAKHQSSAASDPSSSSRRSSQNDAAATAAG-----GTVVMS----- 600
QY 146 PEGKDFVFTGGRGAGHAMVTVASDITEARORILELPEKKGESKAGESKVGELRE 205
DB 601 -----GTRHGQVQRAQPT-SKQATISLLOPPSYKPSNTTQIAQITPLPN 645
QY 206 SNSGANTTETOTSTSSLR--SDPKLMALGT-----VATGLIG----- 244
DB 646 RNSTA-TSSAQDPSTGTGRKADPKRIPLNSTAVQGHRTATGVAANNQGISPHRDH 704
QY 245 -----LATGIVQALALTPEDSPFTTDPDAASATETATDQLTKEAFQ----- 289
DB 705 AQQQVQMNQLTSTYMSKILNKTPAGCTPAATSSSSSSATSTPA--PLQSGSQISHAP 761
QY 290 -----NPDQKXNIDELGNATIS-----GVLDDPVANI--EEOAKAGEA 329
DB 762 TEPVREDDENNSENQNGNVPILIGVGQTSBPVAVQPTEDATSSDKQOQKASSTEP 821
QY 330 KQQAENNAQAQKYDEQQAQKQEBLKVSSGAGYGLSGALLIGGIGVAVTAAHKNOP 389
DB 822 KESKSMIHQSPSPMSQWMTMESIKLSSESGQT--GPTVATGCGPQKATIS-----QQM 874
QY 390 VEQTITTTTTTTTSARTYENKPAANTPAQGVNDTPGSEDTMESRRSSMASTSTFEDTS 449
DB 875 SRSATTSNANMGASSGAAASATQLSGAPSTGASQGYHPKAPSSSSSSST----- 929
QY 450 SIGTVQNP-----YADVKTSLHDSQVPTNSNTSVONMGTDSVVYSTIQHPPR--- 498
DB 930 -----NPPHQHQLTHNHSFSVTPSSYQIPTIS--TAV-NVTSTGPTSSSSAPPRNTR 979
QY 499 -----DTTNGARLLGNPSAG-----IGSTYARL----- 522
DB 980 NRQTFHGTREKDKGDDSDDELGETPGNVISIGATGSSANNAEATIMSKLTLTRDHRE 1039
QY 523 -----ALSGGLRHDHMGGLTGSSNSAVNTSNPPAPGS 554
DB 1040 SMTQFVSGRAGTIGASQGGQTAAALAIAREGGPIAPGA 1078

RESULT 9
A:3412
Hypothetical protein PA1874 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C:Accession: AB3412
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kae, A.; Lapid, K.; Lam,
J.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: AB2950; MUID:2043737; PMID:10984043
A:Accession: AB3412
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-2468 <STD>
A:Cross-references: UNIPROT:Q912M3; UNIPARC:UPI00000054E3; GB:AE004613; GB:AE004091; NID

A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA1874
Query Match 5.9%; Score 167; DB 2; Length 2468;
Best Local Similarity 24.5%; Pred. No. 0.43;
Matches 151; Conservative 53; Mismatches 250; Indels 162; Gaps 34;
QY 2 PIGNLGNVNVNNSIPAPPLPSQTDAGRGQLINSTGSRALFTPVNSMADSGDN 61
DB 701 PIGQVYADSGNWSFPTPTPLPN-----GTVVATATDAS-----GNTSAGSSAT 745
QY 62 RASDVGLPV-NPKMLAASEITLNDGEVFLHDHGPLDITNROIGSSVPRVETOEDKHIA 120
DB 746 VDSVAPATVIND--SNQTLTSTAEF-----GSSV-----TLTDGNGP 783
QY 121 VQGRN-----GVETSVVLSDOEVARLOSIDPEGRKDFVFTGGRGCA--GHAMTVAS 170
DB 784 IGQVYADSGNWSFPTPTPLADGTVVNATATDPAGN-----TSQGSTTVDVGAPTTPVY 838
QY 171 DITTEARQRILELPEKKGESKGA-----GESKVGELRESNSGANTTETOTSTSSSL 225
DB 839 NLSNGSSL-----SGTAEPSGTVILLTDGNGNPIAEVYADSGG--NMTYT----- 880
QY 226 RSDPKLMALGTVATGLIGLATGIVQALALTPEDSP-----TTTPDAAASAT 275
DB 881 ---PSTPIANGTVNVVADAAAGNSPGASVTVDQAPAPVNPBNGTLLSGTAEPPAT 937
QY 276 ETATRDQLTKEAFQNPQNDQKNIIDELGN--ALPSGVLKDDVANIIEQAKAAGEAKQQA 333
DB 938 VTLT-----DGNQNPICQ-VYADSGNWSFPTPTPLANGTVN-----ATASDPT 981
QY 334 INNNAQAQKYDEQQAQKQEBLKVSSGA--GYGLSGALI-LGGIGVAVTAAHKNOP 389
DB 982 GNTSAPASTVVD-SVAPAPVNPBNGAISGTAERGAATVTLTDGSG-----NP 1029
QY 390 VEQTITTTTTTTTSARTYENK--KPAANTPAQGVNDTPGSEDTMESRRSSMASTSTFEDTS 449
DB 1030 IGQVYADSGNWSFPTPTPLADGTVVNATATDPAGNTGQSGSTVDALAPATPTVNLISNG 1089
QY 436 SSMASTS--STFPDTSIGTVQNPVADYKTS-----LHDSQVPTNS--TTSVONMGN 484
DB 1090 SSLSGTAERGSTVILLDNG--NPIAEVYADSGNWTTPSTPIANGTVNVVADDAAG 1146
QY 485 TDS-----VYSTIGHPPRDITDNGARLLGNPSAGIQSTYARLALSGLRHDMGGLTG-G 538
DB 1147 NSSPAPATVVDSSAPPAPVNPBNGVVISGTAEGATVT--LTDAGG--NPIGQVYADG 1201
QY 539 SMSAVNTSNPPAPGS 554
DB 1202 SGNWSFPTPTPLANGT 1217

RESULT 10
T21460
Hypothetical proteine ZK945.10 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T21460
R:Wilkinson, J.
Submitted to the EMBL Data Library, March 1995
A:Reference number: Z19425
A:Accession: T21460
A:Status: Preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-796 <WLL>
A:Cross-references: UNIPARC:UPI000017BD37; EMBL:Z48582; PIRN:CAA88465.1; GSPDB:GN000020;
A:Experimental source: clone F2E5
C:Genetics:
A:Gene: CESP:ZK945.10
A:Map position: 2
A:Introns: 62/3; 124/3; 169/1; 216/1; 262/3; 371/1; 581/3; 608/1

```

Query Match      5.8%; Score 166; DB 2; Length 796;
Best Local Similarity 20.3%; Pred. No. 0.11;
Matches 108; Conservative 89; Mismatches 223; Indels 112; Gaps 19;

Qy 68 GLPNPRLMASETLLND-----GPEVLHDHGPLDTLNROGSSVFRVETDEDKHI 119
   || : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 137 GLPLNSTWITILNENVDDEISIAVEAKYEVCIYDGG----IDRCGSLMW--LÖVGGENM 169

Qy 120 A-VGÖRNGVETSVALSDOEYARLOSIDPEGKDFVTGGRGAGHAMVYASDITTEARÖR 178
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 190 ALLGYRECEBGEI--NEEYARBMCKRPYRSEK-----STAISDQGV 230

Qy 179 ILELLPEKGTCESGAGSGKGVGELRESNGSAEMTTEQSTSTSSLSRDPKMLALGTV 238
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 231 YYDQVLKGRAPAKQPSMKTSGSPFLRRMKRDAGNDCDYLTSTSTSTTTPPTTVTSTV 290

Qy 239 -ATGLIGLAATGIVQALALATPEPDSPTT-----DPDAASATETATRDÖLTKEAFQ 289
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 291 TSTTTVPSTSTVTYAMASTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTST 350

Qy 290 NPDÖKÖVNIDELGNALPESGLKD--DVANIEÖQAKAAG--EAAQÖALENNÖAQÖKRYDE 346
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 351 SPSS-----TLTSLISIPPTTPTTPTTSTLSSLPDNALCSYLDTETTSTPTFTTMTLSTTTE 405

Qy 347 QÖAQKÖEELKXSSGAGYVLSGALLGCGIGAVNLAALHRKÖQÖPEQÖTTTTTTTTT--S 404
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 406 EPGSTTTTEVTI-----TSTVTTTTEPTTTLTSTASTSTEPS 445

Qy 405 ARTYENKPA-----NNTPAÖGNVDTPGSEDTMESRRSSMASTSTPFDTSSIGTVQ 455
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 446 TSTVTTSPSTSPVSTVSTSSSSSTVTTPTSTSTSTSPSTSTSTSTSTSTSTSTSTST 505

Qy 456 NPYADVKTSLSDÖVPTNSNTSVÖMKGNTDSVYRSTIQHPRTÖTNGARLLGNP---511
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 506 S-----STPSSSTAASSVSSSTAOSTGOSTTQÖSTTTKSETTSSSGT---NPDFFYF 555

Qy 512 -----SAGIOGTYARLALSGGLRHDMGGLTGCSSNAVNTSNPPAPGSHREY 558
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 556 VEKATTTTYDSTSVNLNSGL-----GIIYQÖISIECTS-----PISSNYV 597

RESULT 11
B90835
Probable tail fiber protein [imported] - Escherichia coli (strain O157:H7, substrain RIM
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C:Accession: B90835
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
Gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genc
A:Reference number: A59629; MUID:21156231; PMID:11258796
A:Accession: B90835
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-971 <HAY>
A:Cross-references: UNIPROT:Q8XQ04; UNIPARC:UPI000016542B; GB:BA000007; RIND:BA835073.1
A:Experimental source: strain O157:H7, substrain RIMD 0509952
C:Genetics:
A:Gene: ECe150

Query Match      5.8%; Score 164; DB 2; Length 971;
Best Local Similarity 21.2%; Pred. No. 0.18;
Matches 95; Conservative 66; Mismatches 199; Indels 90; Gaps 15;

Qy 79 SEITLNDPEVLH-----DHGPLDTLNROGSSVFRVETÖDEKHIAGÖ-----123
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 59 SVILLVEGFPSPSHAGTIVYEDSGP-GTLNDFLAGM-----TEDDVREALRRFELWEE 112

Qy 124 --RNCVETSVLSDOEFYARLOSIDPEGKDFVTGGRGAGHAMVYASDITTEARÖRILE 181
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 113 VARN---ASAVAKONTAAKKSASDAS-----TSARAANTHA--TDAAADARA-----154

```

```

QY 182 LLEPKGTGSKAGKSGKYGELRNSNGANTETQSTSSLRSPKMLALGTVATG 241
Db 155 -----ASTSAGQAASSAQSSAGTASTATKSKAPAAABSK-----SAAAT 199
QY 242 LIGLATGIVQALALPEPDSPTTTPDDAASATETATRDOLTAEAFQPNQKYNIDEL 301
Db 200 SAGAAKTSTTNAAVSGQSAATSTASTATTKASPAASSARDASAKDEAAKSSFTSAAS----- 255
QY 302 GNAIPSGVLKDDVANIIEQAKAAGEBAK-QQALENNAQAOKKYDEQQAQKOEELKVSSG 360
Db 256 -----SASSAASATTAAGNSAKAAKTSETNAKSETTAAEQSAAAGSK--TA 301
QY 361 AGYGLSGALLIGGCIWAVTAALHRKNQPVQQTITTTTTTTSARFVENKPAANTPAQ- 419
Db 302 AALSSAASATSGAQASASATPA-----GKASESAASSASTATTKGAETEQAASAASSASA 357
QY 420 ---GNVDPRGSEDPYMESRRSSMASTSTFPDTSIGTVONPYADVKTSLHDSQVPTSMN 476
Db 358 AKTSETNAAKASETSASSKPTAAASSASASASASASASASDEATVQAASAAKSSATTASTK 417
QY 477 TSVQNGKNTDSVYVSTIQHPPRDTTNGA 505
Db 418 A-----TEAAGSATPAQAQSKSTAESAA 439

RESULT 12
C85693
probable membrane protein of prophage CP_933X Z1918 [imported] - Escherichia coli (strain
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_rev1sion 16-Feb-2001 #text_change 09-Jul-2004
C:Accession: C85693
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
Miller, L.; Goehreck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: C85693
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-973 <STO>
A:Cross-References: UNIPROT:Q8XDQ4; UNIPARC:UPI000000D0328; GB:AE005174; NID:G12514847; E
A:Experimental source: strain O157:H7, substrain EBL933
C:Genetics:
A:Gene: Z1918

Query Match 5.8%; Score 164; DB 2; Length 973;
Best local Similarity 21.2%; Pred. No. 0.18;
Matches 95; Conservative 66; Mismatches 198; Indels 90; Gaps 15;

QY 79 SEITLNDGFEVL-----DHGPLDTLNKQISSVFRVETOEDGKHIAVGQ----- 123
Db 61 SVILLVEGPPPSHAGITTYVEDSQ-GLNDPLGAM-----TEDDVREALRRFLMYEE 114
QY 124 --RNGVETSVLSDOEVARLQSIDPEGDKRFVFTGSGGAGHAMVTVASDITEARQILIE 181
Db 115 VARN--ASAVAQNTAAAKKSASDAS-----TSAREAAETHA--TDADAARA----- 156
QY 182 LLEPKGTGSKGKSGKYGELRNSNGANTETQSTSSLRSPKMLALGTVATG 241
Db 157 -----ASTSAGQAASSAQSSAGTASTATKSKAPAAABSK-----SAAAT 201
QY 242 LIGLATGIVQALALPEPDSPTTTPDDAASATETATRDOLTAEAFQPNQKYNIDEL 301
Db 202 SAGAAKTSTTNAAVSGQSAATSTASTATTKASPAASSARDASAKDEAAKSSFTSAAS----- 257
QY 302 GNAIPSGVLKDDVANIIEQAKAAGEBAK-QQALENNAQAOKKYDEQQAQKOEELKVSSG 360
Db 258 -----SASSAASATTAAGNSAKAAKTSETNAKSETTAAEQSAAAGSK--TA 303
QY 361 AGYGLSGALLIGGCIWAVTAALHRKNQPVQQTITTTTTTTSARFVENKPAANTPAQ- 419
Db 304 AALSSAASATSGAQASASATPA-----GKASESAASSASTATTKGAETEQAASAASSASA 359

```

Qy 420 ---GNVDTPGSEBDMESRRSSMASTSTPEPTSSIGTVQNPVADVTKSLHDSQVPTNSNV 476
Db 360 AKTSETNAVASERTSASSSTXTPAAASSASASSASASASADEATRQASAKKSATTASRK 419
Qy 477 TSVQNMGNITDSVVYSTIQHPPRDTTNGA 505
Db 420 A-----TEAAGSATMAAQSKSTAESAA 441

RESULT 13

hypothetical protein H02F09.3 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C:Accession: T33369
R:Geisel, C.; Harmon, G.
Submitted to the EMBL Data Library, July 1998
A:Description: The sequence of *C. elegans* cosmid H02F09.
A:Reference number: 221330
A:Accession: T33369
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1275 <GEI>
A:Cross-references: UNIPROT:O76602; UNIPARC:UPI000082AAD; EMBL:AF077538; PIDN:AAC64622.
A:Experimental source: strain Bristol N2; clone H02F09
C:Genetics:
A:Gene: CESP:H02F09.3
A:Map position: X
A:Introns: 42/1, 78/3, 106/3, 135/2, 161/2, 224/1
C:Superfamily: yeast glucan 1,4-alpha-glucosidase homolog; glucan 1,4-alpha-glucosidase

Query Match 5.8%; Score 164; DB 2; Length 1275;

Best Local Similarity 19.6%; Pred. No. 0.26; Matches 111; Conservative 72; Mismatches 235; Indels 148; Gaps 18;

Qy 26 TDGAGRGQLINSTGPIGS-----RALFTFRNSMADSCDNRASDPGLPVNPM 74
Db 764 TDGSGTVSG-----STGSGSTNNPGSDSTTGISTVSGSSLSTISGSGSTVSG----- 813
Qy 75 RLAAEELTNDGEVLHDP-----LDTLNQLGSSVFRVETQEGKHIA----- 120
Db 814 ---SSDMVTVSG-----STSSPGSTESTVSGASTMSPSTGS---VETSTSGSSVSTVSGS 863
Qy 121 ---VGRNGVETSVVLSDE-----YARLQSI 144
Db 864 TSSSTTGOSTVSESSSVSTVSSSTISQSTTGTGSETVFGSTATGSSSTMASSTGST 923
Qy 145 DPEGKDFVFTGGR-----CGAGHAMVTVASDITEARQILLELPKGTESKGA--- 194
Db 924 DTPGTESTITTSVTGSETVSGSTGTTEGSTTISESTMTTV-----GVSTGSTITT 975
Qy 195 GESKGVGELRESNSGAENTETQTSTSTSLRSDPKLMALGTVAITGLATGIVQAL 254
Db 976 GESTVSGSTRSTVTVGSETVSGSTESTSTP-----TVPSTVSGSTGTVTGESVTS 1031
Qy 255 ALTPEDPPTTDPDAASATETATRDQLTKAFQNPDNQV---NIDELNALPSGV 310
Db 1032 GSTASTSGSTGSSSTGASVSGSSASTVSTSGSTGSESTVSGSTVSGSTGSGSTTT 1091
Qy 311 KDDVVAANIEQAKAGEAKQOAIENNAQAKYBQQAQBELKVSAGAYGSGALI 370
Db 1092 GESTVSGSTESTVTVLESTVSGSSVST-----VSANTGSTITTGEST 1131
Qy 371 LGGGIG-----VAVTALARRKNQPVETGTTTTTTTTTTSARVYENKPPANTPAQG 420
Db 1132 VSGSGTSGESTSTILESSTVSVSTGSTRITDGTSSRSRSVTVSASTESTVSGSSASIG 1191
Qy 421 NVDTGSGEDTMSRRSSMASTSTFPDTSIGTVQNPVADVTKSLH-DSQVPTNSNNTSV 479
Db 1192 STNTDSTESTSTGSGTISGSGSTSSSTMSAGT-----GSTETSTSGGTVSGSSLSST- 1245
Qy 480 QNMGNTDSVVYSTIQHPPRDTTNGA 505

Db 1246 ---STRESSSSSTQPPSTSTELLGA 1267

RESULT 14

hypothetical protein pmcC - *Chlamydia trachomatis* (serotype D, strain UW3/Cx)
C:Species: *Chlamydia trachomatis*
C:Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 09-Jul-2004
C:Accession: A71517
R:Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell Science 282, 754-759, 1998
A:Title: Genome sequence of an obligate intracellular pathogen of humans: *Chlamydia tra*
A:Reference number: A71570; MUID:99000809; PMID:9784136
A:Accession: A71517
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1770 <ARN>
A:Cross-references: UNIPROT:O84419; UNIPARC:UPI000047BD1; GB:AE001315; GB:AE001273; NLI
A:Experimental source: serotype D, strain UW-3/Cx
C:Genetics:
A:Gene: pmcC

Query Match 5.8%; Score 164; DB 2; Length 1770;

Best Local Similarity 22.8%; Pred. No. 0.4; Matches 120; Conservative 58; Mismatches 197; Indels 152; Gaps 24;

Qy 118 HIAVGORGVETSVVLSDEVARLQSIDPEGKDFVFTGRCGAGHAMVTVASDITEARQ 177
Db 333 NIATSGAGVFTKENISCTNTNSLOFLKN-----SAGHGCG-AYVTQTMSTVNTTS 383
Qy 178 RLLELEPKGTGE-----SKGAGESKGVGELRNSGAENTETQTSTSTSLRSDPK 230
Db 384 E-SITTPPLIGEVIFSENTAKHGCGICTNLSISN-----LKTIVTLTKNSAKESG-- 433
Qy 231 IMLALGTATGLIGLAATGIVQALATPEPDSPPTTDP-----D 269
Db 434 ---GALFTDLASIPITD-----TPSSSTPSSSPASTREVVASAKINFFASTAKP 481
Qy 270 AAASATETATRDQI--TKEAFQNPDNQKVINIDELNALPSGVLKDDVVAANIEQAKAA-- 325
Db 482 AAPSLTE-AESDQTQETSTSDTSDID-VSIEINILN-----VAINQNTSAKKGA 529
Qy 326 ---GEBAKQOAIENNAQAKYBQQAQBELKVSAGAYGSGALIIG-----GGTGYA 378
Db 530 IYKKAQKLSRINN-----LELGSNSQDVGGGLCTLESVEFPALIGL 571
Qy 379 VT-----AALRRKNQPVETGTTTTTTTTTTSARVYEN---KPPANTPAQGNVDTPG 426
Db 572 LSHVNSAAKEGALHSKTVTLNLSKSTTFPADNTKAIVESTPEAPDEELPVEGEBST-A 630
Qy 427 SEDTMSRRSSMASTSTSTPEPTSSIGT--VONPYADV-----K 462
Db 631 TEDPNSNTEGSSANTNLBSGSGDTADTGTCPDNNSQDPTSDGNABERQLDSTQSNNE 690
Qy 463 TSLHDSQVPTNSNNTSVQNMGNITDSV---VYSTIQHPPRDTTNGARILNPS----- 512
Db 691 NTLPNNSNIDOSNENNTDESSTHEITDESVSSSSSGSSTPQDGAASGAPSGDOSIS 750
Qy 513 ---AGISTYARLALSGALRHDMCG---LTGGSNSAVNTSNPNPAPGS 554
Db 751 ANACIAKSYAASTDSPPVSNSSGSEEPVTSSTSDSDVTYASSDNPDS 797

RESULT 15

protein H39E23.1 (imported) - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
C:Accession: G89287
R:anonymus, the C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode *C. elegans*: a platform for investigating biolog
A:Reference number: A75000; MUID:99069613; PMID:9851916

A:Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_el
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A:Accession: G99287
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1246 <STO>
A:Cross-references: UNIPARC:UPI000017a692; GB:chr_V; PIDN:CAB09532.1; PID:G3878100; GSPD
C:Genetics:
A:Gene: H39E23.1
A:Map position: 5

```
Query Match          5.8%; Score 163.5; DB 2; Length 1246;  
Best Local Similarity 19.0%; Pred. No. 0.27;  
Matches 126; Conservative 76; Mismatches 242; Indels 219; Gaps 22;  
  
Qy 36 INSTGPIGSRALFTFVRNSMADSGDNRASDVPGLEPVMRLAASEITTLNDGFEVLHDGP 95  
Db 395 INVSSSLQHPAGVITREHVT-----SSSASGSSASPSRYSRSSATAT-GASITAGSAL 447  
Qy 96 LDTLNRO-----IGSVFRVETQEDGKHIAVGQRNGVETSVLSDOEYARLQSID 145  
Db 448 ASANAAQKHQSSAAPSSSSSSRSSQNDAAATAG-----GTVMS----- 490  
Qy 146 PEGDKFVFTGGAGHAMVTVASDITEARQRIELLEPKGTGSKGAGSKGVGLRE 205  
Db 491 -----GTRHGVQWRAQPT-SRQATISLLQPPSYKPSNNTQIAQIPLFN 535  
Qy 206 SNSGAETTEGTSTSTSLR--SDPKMLALGT-----VATGLIG----- 244  
Db 536 RNSTA-TSSAAQPSGTGTRKIDPKGRIPLNSTAVQGHRTATGAVANNGIIPSHRDH 594  
Qy 245 -----LAATGIVQALATPEPDSPTTTPDAASATETATRDQLTKEAFQ----- 289  
Db 595 AQQQQYNNQLTSTTMSKLINKTPAAGTAATSSSSSSSATSTA--PLQKSGSQISHAP 651  
Qy 290 -----NPDNQKVINIDELGNAIPS-----GVLKDDVVANI---EEQAKAGEEA 329  
Db 652 TEPTREDDDENNSNGVNPPLIGVGPOQTSRAVQVPTEDATSSDKKQOQKASSTP 711  
Qy 330 KQ-----QAIENNAQAQKTDQOAKROELKVSAGYGLSGALLIGGGIGAVVTAL 383  
Db 712 KESNPIVQNLHLNSLKLSDSSAATSYPEPRPGIAG----- 750  
Qy 384 HRKNQPVFQTTTTTTTTTSARTVENKPNANTP-----AQ 419  
Db 751 -RRSEPSAATRRRRQTMVNDARKLQTPPDIDRPHFEDTTIDROMALYVSTASSRMTR 809  
Qy 420 GNVDTPGSEDTWESR-----RSSMASTSTFFDTSSIGTV--- 454  
Db 810 GVLPTPTPSNSTSSSFIVEPLTHVAASPDITTTTPTKSTVTTSYPFRRTPSFRVLIVL 869  
Qy 455 -----QNPYADVKTSLHDSQVPTSNSNTSVQNMGNTDSVVYSTIOHPPRD 499  
Db 870 LLDNGRLMPSMIHQSP-----SMPPSQWMTAMESLKLSESQGTGGPVTATGPPORA 923  
Qy 500 TTIDNGARLLGNPSAGIOSITYAARLALSGLRHMGGI---TGGS-----NSAVNTSN 547  
Db 924 TSQOMRSATTTSAMMGASSGGAAAAAASATNOISGAPSTGASSQOYHPKAPSSSSSST 983  
Qy 548 NPP 550  
Db 984 NPP 986
```

Search completed: August 1, 2006, 22:14:04
Job time : 46 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: August 1, 2006, 22:03:00 ; Search time: 303 Seconds

(without alignments)
1703.494 Million cell updates/sec

Title: US-09-189-415d-11

Perfect score: 2840
Sequence: 1 MFIGNLGHNPNVNSIPAP.....SNSAVNTSNNPPAPSGHRFV 558

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2849598 seqs, 925015592 residues

Total number of hits satisfying chosen parameters: 2849598

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Uniprot 7.2:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2840	100.0	558	09R396_ECOLI	09R396 escherichia
2	2840	100.0	558	07DB77_ECO57	07DB77 escherichia
3	2824	99.4	558	085506_ECOLI	085506 escherichia
4	2808	98.9	558	0421M4_ECOLI	0421M4 escherichia
5	2747	96.7	574	058188_ECOLI	058188 escherichia
6	1835	64.9	538	047014_ECOLI	047014 escherichia
7	1835	64.6	538	047016_ECOLI	047016 escherichia
8	1835	64.6	538	05K5P9_ECOLI	05K5P9 escherichia
9	1832	64.5	538	058189_ECOLI	058189 escherichia
10	1824	64.2	538	085508_ECOLI	085508 escherichia
11	1820	64.1	538	058190_ECOLI	058190 escherichia
12	1820	64.1	538	058197_ECOLI	058197 escherichia
13	1613.5	56.8	551	068258_ECOLI	068258 escherichia
14	1611.5	56.7	551	0421M1_ECOLI	0421M1 escherichia
15	1607.5	56.6	551	0421M0_ECOLI	0421M0 escherichia
16	1594	56.1	552	09KMH9_ECOLI	09KMH9 escherichia
17	1585	55.8	550	052147_ECOLI	052147 escherichia
18	1581	55.7	550	05MWC9_ECOLI	05MWC9 escherichia
19	1569.5	55.3	547	07BHL5_GENTR	07BHL5 citrobacter
20	1569.5	55.3	547	09WAK1_ECOLI	09WAK1 escherichia
21	1568.5	55.2	547	09ERT11_GENTR	09ERT11 citrobacter
22	1531.5	53.9	549	050190_ECOLI	050190 escherichia
23	1146.5	40.4	367	079C12_ECOLI	079C12 escherichia
24	215.5	7.6	1374	07YU77_DROME	07YU77 drosophila
25	214.5	7.6	1373	07KSZ0_DROME	07KSZ0 drosophila
26	214.5	7.6	1376	09V163_DROME	09V163 drosophila
27	208.5	7.3	1321	0962D1_DROME	0962D1 drosophila
28	207.5	7.0	1315	0882W9_DROME	0882W9 drosophila
29	200	6.9	2332	081FX6_CABEL	081FX6 caenorhabditis
30	196	6.9	2338	0759S3_ASHGO	0759S3 ashbya goss
31	195.5	6.9	1254	094185_CABEL	094185 caenorhabditis

32	190.5	6.7	1466	2	04PFY2_USITMA	04PFY2 usitlago ma
33	188	6.6	1184	2	041HX0_GIBZE	041HX0 gibberella
34	187	6.6	1281	2	077261_DROME	077261 drosophila
35	187	6.6	1300	2	09W5E0_DROME	09W5E0 drosophila
36	185	6.5	2310	2	08CMU7_STABE	08CMU7 staphylococcus
37	184.5	6.5	3135	2	07KUH2_DROME	07KUH2 drosophila
38	183	6.4	860	2	05CTC7_CRYPV	05CTC7 cryptosporidium
39	182	6.4	574	2	07RW61_NEUCR	07RW61 neurospora
40	182	6.4	3080	2	09V602_DROME	09V602 drosophila
41	182	6.4	3109	2	09BWO0_DROME	09BWO0 drosophila
42	181	6.4	626	2	05SGA0_HUMAN	05SGA0 homo sapien
43	180	6.3	596	2	06UXC5_HUMAN	06UXC5 homo sapien
44	180	6.3	1086	2	059X10_CANAL	059X10 candida alb
45	179.5	6.3	901	2	04W9Q8_ASEFV	04W9Q8 aspergillus

ALIGNMENTS

RESULT 1
09R396_ECOLI PRELIMINARY; PRT; 558 AA.
ID 09R396; AC 09R396;
DT 01-MAY-2000, integrated into UniprotKB/TREMBL.
DT 01-MAY-2000, sequence version 1.
DT 07-FEB-2006, entry version 25.
DE Translocated into/in receptor TIR (L00277).
GN Name=tlr;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN (1)
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=86/24;
RX MEDLINE=99242825; PubMed=10225900;
RA Devaney R., Stein M., Reinecheld D., Abe A., Ruskowski S.,
RA Finlay B.B.;
RT "Enterohemorrhagic Escherichia coli O157:H7 produces TIR, which is
RT translocated to the host cell membrane but is not tyrosine
RT phosphorylated.";
RL Infect. Immun. 67:2389-2398(1999).
RN (2)
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC43895, and EDL933;
RX MEDLINE=98339885; PubMed=9673266;
RA Perna N.T., Mayhew G.F., Poefel G., Elliott S., Donnenberg M.S.,
RA Kaper J.B., Blatterner F.R.;
RT "Molecular evolution of a pathogenicity island from enterohemorrhagic
RT Escherichia coli O157:H7.";
RL Infect. Immun. 66:3810-3817(1998).
RN (3)
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CP97;
RX PubMed=16272509; DOI=10.1128/JCM.43.11.5715-5720.2005;
RA Garandia J., Ren Z., Tennant S., Midoll Viera M.A., Chong Y.,
RA Whale A., Azopardo K., Dahan S., Siricili M.P., Franzolin M.R.,
RA Trubets L.R., Phillips A., Gomes T.A., Xu J., Robins-Browne R.,
RA Frankel G.;
RT "Distribution of tccp in clinical Enterohemorrhagic and
RT Enteropathogenic Escherichia coli Isolates.";
RL J. Clin. Microbiol. 43:5715-5720(2005).
CC Copyrighted by the Uniprot Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonCommercial License
CC
CC EMBL: AF125993; AAD29391.1; -; Genomic DNA.
CC EMBL: AF071034; AAC31506.1; -; Genomic DNA.
CC EMBL: DQ007021; AAY25392.1; -; Genomic DNA.
CC PIR: A98199; A98199.
CC PIR: E86045; E86045.
CC HSP: 09KMH9; 1P02.
CC SMR: 09R396; 269-333.

DR BiOCyc; ECOl8334-1:ECs4561-MONOMER; -;
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR003536; Ttr_rcpt.
DR Pfam; PF07489; Ttr_receptor_C_1.
DR Pfam; PF03549; Ttr_receptor_M_1.
DR Pfam; PF07490; Ttr_receptor_N_1.
DR PRINTS; PR01370; TRANSINTIMINR.
KM Receptor.
SQ SEQUENCE 558 AA; 58022 MW; 99C417222D4B4A41 CRC64;

Query Match 100.0%; Score 2840; DB 2; Length 558;
Best Local Similarity 100.0%; Pred. No. 2,1e-142;
Matches 558; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPIGNLGNPNVNNISIPAPPLPSQTDGAGRGQLINSTGRLGSLFTPVNNSMADSGD 60
DB 1 MPIGNLGNPNVNNISIPAPPLPSQTDGAGRGQLINSTGRLGSLFTPVNNSMADSGD 60

QY 61 NRASDVPLPVNPMKLAASEITLNDGFEVLHDHGLDITLRQIGSSVFRVETQEDGKHIA 120
DB 61 NRASDVPLPVNPMKLAASEITLNDGFEVLHDHGLDITLRQIGSSVFRVETQEDGKHIA 120

QY 121 VQQRNGVETSVVLSDOEYARLQSIDPEGKDFVFTGRCGAGHAMVTVASDITEARQRL 180
DB 121 VQQRNGVETSVVLSDOEYARLQSIDPEGKDFVFTGRCGAGHAMVTVASDITEARQRL 180

QY 181 ELLEPKGTGSGKAGESKGVGLRESNSGAENTTETQSTSTSLRSDPKMLALGTAT 240
DB 181 ELLEPKGTGSGKAGESKGVGLRESNSGAENTTETQSTSTSLRSDPKMLALGTAT 240

QY 241 GLIGLAATGIQVALALTPEDPSPTTTPDPAASATEIATRDQITKEAFONPNQKYNIDE 300
DB 241 GLIGLAATGIQVALALTPEDPSPTTTPDPAASATEIATRDQITKEAFONPNQKYNIDE 300

QY 241 GLIGLAATGIQVALALTPEDPSPTTTPDPAASATEIATRDQITKEAFONPNQKYNIDE 300
DB 241 GLIGLAATGIQVALALTPEDPSPTTTPDPAASATEIATRDQITKEAFONPNQKYNIDE 300

QY 301 LGNAIPSGVLKDDVVAANIEBQAKAAGEAKQQAIEENNAQAKKYDEBQAKQEBLKVS 360
DB 301 LGNAIPSGVLKDDVVAANIEBQAKAAGEAKQQAIEENNAQAKKYDEBQAKQEBLKVS 360

QY 361 AGYGSGLIILGGIGVAVTALHRKNQVVEQTITTTTTTTSARVENKKNANTPAQG 420
DB 361 AGYGSGLIILGGIGVAVTALHRKNQVVEQTITTTTTTTSARVENKKNANTPAQG 420

QY 421 NVDTGSESDTMRSSMASTSTSTFTDSSIGTVONPYADVKTSLHDSOVPTSNSTSVQ 480
DB 421 NVDTGSESDTMRSSMASTSTSTFTDSSIGTVONPYADVKTSLHDSOVPTSNSTSVQ 480

QY 481 NMGNNTDSVYVSTIQHPPRDTTNGARLLGNPSAGIQSTYARLALSGGLRHDWGLTGGSN 540
DB 481 NMGNNTDSVYVSTIQHPPRDTTNGARLLGNPSAGIQSTYARLALSGGLRHDWGLTGGSN 540

QY 541 SAVNTSNNPAPGSHRFV 558
DB 541 SAVNTSNNPAPGSHRFV 558

RESULT 2
Q7DB77_ECO57 PRELIMINARY; PRT; 558 AA.
AC Q7DB77; Q7A9Q1;
DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.
DT 05-JUL-2004, sequence version 1.
DT 07-FEB-2006, entry version 17.
DE Puctive translocated intimin receptor protein (Translocated intimin
receptor Tlr).
GN Name=tlr; OrderedlocusName=ECs4561, z5112;
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxId=83334;
RN (1)
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=O157:H7 / EDL933 / ATCC 700927 / EHEC;

RX MEDLINE=21074935; PubMed=11206551; DOI=10.1038/35054089;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glaesner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamoukis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533(2001).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=O157:H7 / Sakai / RIMD 0509952 / EHEC;
RX MEDLINE=21156231; PubMed=11258796; DOI=10.1093/dnares/8.1.11;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kubara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NonDerivs license
CC -----
CC EMBL; AE005174; AAG58825.1; -; Genomic DNA.
CC EMBL; BA000007; BAB37984.1; -; Genomic DNA.
CC SMR; Q7DB77; 269-333.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR003536; Ttr_rcpt.
DR Pfam; PF07489; Ttr_receptor_C_1.
DR Pfam; PF03549; Ttr_receptor_M_1.
DR Pfam; PF07490; Ttr_receptor_N_1.
DR PRINTS; PR01370; TRANSINTIMINR.
KM Complete proteome; Receptor.
SQ SEQUENCE 558 AA; 58022 MW; 99C417222D4B4A41 CRC64;

Query Match 100.0%; Score 2840; DB 2; Length 558;
Best Local Similarity 100.0%; Pred. No. 2,1e-142;
Matches 558; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPIGNLGNPNVNNISIPAPPLPSQTDGAGRGQLINSTGRLGSLFTPVNNSMADSGD 60
DB 1 MPIGNLGNPNVNNISIPAPPLPSQTDGAGRGQLINSTGRLGSLFTPVNNSMADSGD 60

QY 61 NRASDVPLPVNPMKLAASEITLNDGFEVLHDHGLDITLRQIGSSVFRVETQEDGKHIA 120
DB 61 NRASDVPLPVNPMKLAASEITLNDGFEVLHDHGLDITLRQIGSSVFRVETQEDGKHIA 120

QY 121 VQQRNGVETSVVLSDOEYARLQSIDPEGKDFVFTGRCGAGHAMVTVASDITEARQRL 180
DB 121 VQQRNGVETSVVLSDOEYARLQSIDPEGKDFVFTGRCGAGHAMVTVASDITEARQRL 180

QY 181 ELLEPKGTGSGKAGESKGVGLRESNSGAENTTETQSTSTSLRSDPKMLALGTAT 240
DB 181 ELLEPKGTGSGKAGESKGVGLRESNSGAENTTETQSTSTSLRSDPKMLALGTAT 240

QY 241 GLIGLAATGIQVALALTPEDPSPTTTPDPAASATEIATRDQITKEAFONPNQKYNIDE 300
DB 241 GLIGLAATGIQVALALTPEDPSPTTTPDPAASATEIATRDQITKEAFONPNQKYNIDE 300

QY 301 LGNAIPSGVLKDDVVAANIEBQAKAAGEAKQQAIEENNAQAKKYDEBQAKQEBLKVS 360
DB 301 LGNAIPSGVLKDDVVAANIEBQAKAAGEAKQQAIEENNAQAKKYDEBQAKQEBLKVS 360

QY 361 AGYGSGLIILGGIGVAVTALHRKNQVVEQTITTTTTTTSARVENKKNANTPAQG 420
DB 361 AGYGSGLIILGGIGVAVTALHRKNQVVEQTITTTTTTTSARVENKKNANTPAQG 420

QY 421 NVDTGSESDTMRSSMASTSTSTFTDSSIGTVONPYADVKTSLHDSOVPTSNSTSVQ 480
DB 421 NVDTGSESDTMRSSMASTSTSTFTDSSIGTVONPYADVKTSLHDSOVPTSNSTSVQ 480

QY 481 NMGNTDSVYVSTIQHPRTDNGARLLGNPSAGIOSTYARLALSGGLRHDGGLTGSGN 540
DB 481 NMGNTDSVYVSTIQHPRTDNGARLLGNPSAGIOSTYARLALSGGLRHDGGLTGSGN 540
QY 541 SAVNTSNPNPAPGSHRFV 558
DB 541 SAVNTSNPNPAPGSHRFV 558

RESULT 3
ID 085506_ECOLI PRELIMINARY; PRT; 558 AA.
AC 085506;
DT 01-NOV-1998, integrated into UniProtKB/TrEMBL.
DT 01-NOV-1998, sequence version 1.
DT 07-FEB-2006, entry version 18.
DE Translocated intilmin receptor TIR.
GN Name=tir;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxId=562;
RN (1)
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=95SF2;
RX MEDLINE=99003184; PubMed=9784578;
RA Paton A.W., Manning P.A., Woodrow M.C., Paton J.C.;
RT "Translocated intilmin receptors (TIR) of Shiga-toxicogenic Escherichia
RT coli isolates belonging to serogroups O26, O111, and O157 react with
RT sera from patients with hemolytic-uremic syndrome and exhibit marked
RT sequence heterogeneity.";
RT Infect. Immun. 66:5580-5586(1998).
CC
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NonDerivs license

DR EMBL: AF070067; AAC69314.1; -; Genomic_DNA.
DR HSSP: O9KM9; 1F02.
DR GO: GO:0004872; F:receptor activity; IEA.
DR GO: GO:0007155; P:cell adhesion; IEA.
DR InterPro: IPR003536; TIR_rcpt.
DR Pfam: PF07489; TIR_receptor_C; 1.
DR Pfam: PF03549; TIR_receptor_M; 1.
DR Pfam: PF07490; TIR_receptor_N; 1.
DR PRINTS: PR01370; TRANSINTIMINR.
KM Receptor.
SQ SEQUENCE 558 AA; 58176 MW; CA2CDDA94527C2E CRC64;

Query Match 99.4%; Score 2824; DB 2; Length 558;
Best Local Similarity 99.5%; Pred. No. 1.5e-141;
Matches 555; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MPIGNLGNPNVNNISIPAPPLPSQTDGAGRGQLINSTGPLGRALFTPVNSMADSGD 60
DB 1 MPIGNLGNPNVNNISIPAPPLPSQTDGAGRGQLINSTGPLGRALFTPVNSMADSGD 60
QY 61 NRASDVPGLPVNPMLAASEITLNDGFEVLHDHGPLDTLNRQIGSSVFRVETQEDGKHIA 120
DB 61 NRASDVPGLPVNPMLAASEITLNDGFEVLHDHGPLDTLNRQIGSSVFRVETQEDGKHIA 120
QY 121 VQORNGVETSVLSQOEVARLQSIDPEGKDFVFTGCRGAGAHAMTVASDITTEAQRLL 180
DB 121 VQORNGVETSVLSQOEVARLQSIDPEGKDFVFTGCRGAGAHAMTVASDITTEAQRLL 180
QY 181 ELLEPKGTGSGKAGSGKVGELRESNGAENTTETQTSTSTSLRSDPKMLALGTVAI 240
DB 181 ELLEPKGTGSGKAGSGKVGELRESNGAENTTETQTSTSTSLRSDPKMLALGTVAI 240
QY 241 GLIGLAATGI VQALALTPEPDSTTTDPAAASATETATRDQITKFAFPNPNQKNYIDE 300
DB 241 GLIGLAATGI VQALALTPEPDSTTTDPAAASATETATRDQITKFAFPNPNQKNYIDE 300
QY 301 LGNAIPSGVLKODVVAНИЕEQAKAGEAKQQAIEENNAQOKKYDEQAKRQBELKVSSG 360

DB 301 LGNAIPSGVLKODVVAНИЕEQAKAGEAKQQAIEENNAQOKKYDEQAKRQBELKVSSG 360
QY 361 AGYGLSGAILLGIGIVAVTAALHRKQVPEQTTTTTTTTTSAITVENKPNANTPAQG 420
DB 361 AGYGLSGAILLGIGIVAVTAALHRKQVPEQTTTTTTTTTSAITVENKPNANTPAQG 420
QY 421 NVDTGSEDTMSRRSSMASTSTFPDTSIGTVQNPYADVKTSLHDSQVPTSNSITVQ 480
DB 421 NVDTGSEDTMSRRSSMASTSTFPDTSIGTVQNPYADVKTSLHDSQVPTSNSITVQ 480
QY 481 NMGNTDSVYVSTIQHPRTDNGARLLGNPSAGIOSTYARLALSGGLRHDGGLTGSGN 540
DB 481 NMGNTDSVYVSTIQHPRTDNGARLLGNPSAGIOSTYARLALSGGLRHDGGLTGSGN 540
QY 541 SAVNTSNPNPAPGSHRFV 558
DB 541 SAVNTSNPNPAPGSHRFV 558

RESULT 4
ID 042IM4_ECOLI PRELIMINARY; PRT; 558 AA.
AC 042IM4;
DT 07-JUN-2005, integrated into UniProtKB/TrEMBL.
DT 07-JUN-2005, sequence version 1.
DT 07-FEB-2006, entry version 5.
DE Translocated intilmin receptor.
GN Name=tir;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxId=562;
RN (1)
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CPG122-G57, and CPG6;
RX PubMed=16272509; DOI=10.1128/JCM.43.11.5715-5720.2005;
RA Garnemdia U., Ren Z., Tennant S., Midolli Viera M.A., Chong Y.,
RA Whale A., Azopardo K., Dahan S., Sircill M.P., Franzolin M.R.,
RA Trublet L.R., Phillips A., Gomes T.A., Xu J., Robins-Browne R.,
RA Frankel G.;
RT "Distribution of tccp in Clinical Enterohemorrhagic and
RT Enteropathogenic Escherichia coli Isolates.";
RT J. Clin. Microbiol. 43:5715-5720(2005).
CC
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NonDerivs license

DR EMBL: DQ007020; AAY25391.1; -; Genomic_DNA.
DR EMBL: DQ007019; AAY25390.1; -; Genomic_DNA.
DR SMR: 042IM4; 271-335.
DR GO: GO:0004872; F:receptor activity; IEA.
DR GO: GO:0007155; P:cell adhesion; IEA.
DR InterPro: IPR003536; TIR_rcpt.
DR Pfam: PF07489; TIR_receptor_C; 1.
DR Pfam: PF03549; TIR_receptor_M; 1.
DR Pfam: PF07490; TIR_receptor_N; 1.
DR PRINTS: PR01370; TRANSINTIMINR.
KM Receptor.
SQ SEQUENCE 558 AA; 58008 MW; 69ELAD5880AFB10 CRC64;

Query Match 98.9%; Score 2808; DB 2; Length 558;
Best Local Similarity 99.3%; Pred. No. 1e-140;
Matches 556; Conservative 0; Mismatches 0; Indels 4; Gaps 2;

QY 1 MPIGNLGNPNVNNISIPAPPLPSQTDGAGRGQLINSTGPLGRALFTPVNSMADSGD 60
DB 1 MPIGNLGNPNVNNISIPAPPLPSQTDGAGRGQLINSTGPLGRALFTPVNSMADSGD 60
QY 61 NRASDVPGLPVNPMLAASEITLNDGFEVLHDHGPLDTLNRQIGSSVFRVETQEDGKHIA 120
DB 61 NRASDVPGLPVNPMLAASEITLNDGFEVLHDHGPLDTLNRQIGSSVFRVETQEDGKHIA 120

QY 121 VQNRNGVETSVVLSDOEYARLQSIDPEKDKFVTGRCGAGHAMVTVASDITEARQIL 180
 DB 121 VQNRNGVETSVVLSDOEYARLQSIDPEKDKFVTGRCGAGHAMVTVASDITEARQIL 180
 QY 181 ELLEPKGTGSGKAGESKGVGLRESNSGAENTTETQ--TSTSTSLRSDPKMLALGTV 238
 DB 181 ELLEPKGTGSGKAGESKGVGLRESNSGAENTTETQ--TSTSTSLRSDPKMLALGTV 240
 QY 229 ATGLIGLAATGIVQALATLPEPDSPTTDDPAAATATATATDQTLTKAFQPNQKNI 298
 DB 241 ATGLIGLAATGIVQALATLPEPDSPTTDDPAAATATATATDQTLTKAFQPNQKNI 300
 QY 299 DELGNAIPSGVLKDDVANIIEQAAKAGEAKQAQIENNAQOKKYDQQAQROBELKYS 358
 DB 301 DELGNAIPSGVLKDDVANIIEQAAKAGEAKQAQIENNAQOKKYDQQAQROBELKYS 360
 QY 359 SGAGYGLSGALILGGGIGVAVTAAALHRKNQVPEQTTTTTTTTTSAATVENKPAANTPA 418
 DB 361 SGAGYGLSGALILGGGIGVAVTAAALHRKNQVPEQ--TTTTTTTTTSAATVENKPAANTPA 418
 QY 419 QGNVDTPESEDTEMRSSMASTSTFPDTSIGTVQNPYADVKTSLHDSQVPTSNST 478
 DB 419 QGNVDTPESEDTEMRSSMASTSTFPDTSIGTVQNPYADVKTSLHDSQVPTSNST 478
 QY 479 VQNMNTDSVYVSTIQHPPRDTTNGARLLGNPSAGIOSTYARLALSGGLRHDGGLTGG 538
 DB 479 VQNMNTDSVYVSTIQHPPRDTTNGARLLGNPSAGIOSTYARLALSGGLRHDGGLTGG 538
 QY 539 SNSAVNTSNPPAPGSHRFV 558
 DB 539 SNSAVNTSNPPAPGSHRFV 558

RESULT 5

Q58188_ECOLI PRELIMINARY; PRT; 574 AA.
 AC Q58188;
 DT 26-APR-2005, integrated into UniProtKB/TrEMBL.
 DT 26-APR-2005, sequence version 1.
 DT 07-FEB-2006, entry version 6.
 DE Translocated intimin receptor.
 GN Name=titr;
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 CX NCBI_TaxID=562;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=Balsman's Run stream;
 RX PubMed:15870341; DOI=10.1128/AEM.71.5.2511-2519.2005;
 RA Higgins J.A., Belt K.T., Karns J.S., Russell-Anelli J., Shelton D.R.;
 RT "titr- and stx-positive Escherichia coli in Stream Waters in a
 Metropolitan Area.";
 RL Appl. Environ. Microbiol. 71:2511-2519(2005).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=Balsman's Run stream;
 RA Hohn C., Karns J.S., Higgins J.A.;
 RL Submitted (FEB-2005) to the EMBL/GenBank/DBJ databases.
 CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
 CC Distributed under the Creative Commons Attribution-NonDerivs license
 CC -----
 CC EMBL: AY944737, AAX47730.1; -; Genomic_DNA.
 DR SMR; Q58188; 287-351.
 DR GO; GO:0004872; F:receptor activity; IEA.
 DR GO; GO:0007155; P:cell adhesion; IEA.
 DR InterPro: IPR003536; TIR rcpt.
 DR Pfam: PF07489; TIR_receptor_C; 1.
 DR Pfam: PF03549; TIR_receptor_M; 1.
 DR Pfam: PF07490; TIR_receptor_N; 1.
 DR PRINTS; PR01370; TRANSINTIMINR.
 KW Receptor.

SEQ SEQUENCE 574 AA; 59314 MW; DC78996D1BA3F5F0 CRC64;
 Query Match 96.7%; Score 2747; DB 2; Length 574;
 Best Local Similarity 95.3%; Pred. No. 1,8e-13;
 Matches 549; Conservative 1; Mismatches 6; Indels 20; Gaps 3;
 QY 1 MPIGNLGNPNVNNISIPAPPLPSQTDAGGRGQILINSTGLSGRALFTPVNSMADSGD 60
 DB 1 MPIGNLGNPNVNNISIPAPPLPSQTDAGGRGQILINSTGLSGRALFTPVNSMADSGD 60
 QY 61 NRASDVPGLPVNPRLLASBITLNDGFEVLHDGPLEDTLNRQIGSSVRFVTEQEDKRIA 120
 DB 61 NRASDVPGLPVNPRLLASBITLNDGFEVLHDGPLEDTLNRQIGSSVRFVTEQEDKRIA 120
 QY 121 VQNRNGVETSVVLSDOEYARLQSIDPEKDKFVTGRCGAGHAMVTVASDITEARQIL 180
 DB 121 VQNRNGVETSVVLSDOEYARLQSIDPEKDKFVTGRCGAGHAMVTVASDITEARQIL 180
 QY 181 ELLEPKGT-----GSKAGESKGVGLRESNSGAENTTETQ-----TSTST 222
 DB 181 ELLEPKGTGSGKAGESKGVGLRESNSGAENTTETQ-----TSTST 222
 QY 223 SLRSDPKMLALGTVATGLIGLAATGIVQALATLPEPDSPTTDDPAAATATATRDQ 282
 DB 241 SLRSDPKMLALGTVATGLIGLAATGIVQALATLPEPDSPTTDDPAAATATATRDQ 300
 QY 283 LTKEAFQPNQKNIINBELGNAIPSGVLKDDVANIIEQAAKAGEAKQAQIENNAQOK 342
 DB 301 LTKEAFQPNQKNIINBELGNAIPSGVLKDDVANIIEQAAKAGEAKQAQIENNAQOK 360
 QY 343 KYDEQAKROBELKYSAGYGLSGALILGGGIGVAVTAAALHRKNQVPEQTTTTTTTTT 402
 DB 361 KYDEQAKROBELKYSAGYGLSGALILGGGIGVAVTAAALHRKNQVPEQ--TTTTTTTT 418
 QY 403 TSARTVENKPAANTPAQGNVDTPESEDTEMRSSMASTSTFPDTSIGTVQNPYADV 462
 DB 419 TSARTVENKPAANTPAQGNVDTPESEDTEMRSSMASTSTFPDTSIGTVQNPYADV 478
 QY 463 TSLHDSQVPTSNSTSVQNMNTDSVYVSTIQHPPRDTTNGARLLGNPSAGIOSTYARL 522
 DB 479 TSLHDSQVPTSNSTSVQNMNTDSVYVSTIQHPPRDTTNGARLLGNPSAGIOSTYARL 538
 QY 523 ALSGGLRHDGGLTGGSNSAVNTSNPPAPGSHRFV 558
 DB 539 ALSGGLRHDGGLTGGSNSAVNTSNPPAPGSHRFV 574

RESULT 6

Q47014_ECOLI PRELIMINARY; PRT; 538 AA.
 AC Q47014;
 DT 01-NOV-1996, integrated into UniProtKB/TrEMBL.
 DT 01-MAY-1999, sequence version 2.
 DT 07-FEB-2006, entry version 24.
 DE Translocated intimin receptor Titr (Translocated intimin co-receptor)
 DE (Beige protein).
 GN Name=titr; Synonyms=espT;
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 CX NCBI_TaxID=562;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=REPEC 84/110/1, and E65/56;
 RA Krejany E.O.;
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=957G1;
 RX MEDLINE=99003184; PubMed=9784578;
 RA Paton A.W., Manning P.A., Woodrow M.C., Paton J.C.;
 RT "Translocated intimin receptors (Titr) of Shiga-toxicogenic Escherichia
 coli isolates belonging to serogroups O26, O111, and O157 react with


```

RA Boedeker E.C.;
RT "Complete nucleotide sequence and analysis of the locus of enterocyte
RT effacement from rabbit diarrheagenic Escherichia coli RDEC-1.";
RL Infect. Immun. 69:2107-2115(2001).
RN
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=83/39;
RX MEDLINE=22063667; PubMed=12067342;
RA Tauschek M., Strugnell R.A., Robins-Browne R.M.;
RT "Characterization and evidence of mobilization of the LEE
RT pathogenicity island of rabbit-specific strains of enteropathogenic
RT Escherichia coli.";
RL Mol. Microbiol. 44:1533-1550(2002).
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs license
CC
CC -----
CC EMBL; U59504; AAD19750.1; -; Genomic_DNA.
CC EMBL; AF045568; AAC15683.1; -; Genomic_DNA.
CC EMBL; AF200363; AAK26722.1; -; Genomic_DNA.
CC EMBL; AF453441; AAL57549.1; -; Genomic_DNA.
CC HSRP; O9KWH9; 1P02.
DR SMR; Q47016; 261-325.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR003536; T1r rcpt.
DR Pfam; PF07489; T1r_receptor_C; 1.
DR Pfam; PF03549; T1r_receptor_M; 1.
DR Pfam; PF07490; T1r_receptor_N; 1.
DR PRINTS; PR01370; TRANSINTIMINR.
KM Receptor.
SQ SEQUENCE 538 AA; 55411 MW; 3132A96987B3D06C CRC64;

Query Match 64.6%; Score 1835; DB 2; Length 538;
Best Local Similarity 65.7%; Pred. No. 3.9e-89;
Matches 370; Conservative 61; Mismatches 102; Indels 30; Gaps 6;

QY 1 MPIGLGNPNVNNISIPAPPLPSQTDGAGG-RGQLINSTGPGLSRALFTPRNSMADSG 59
DB 1 MPIGLGNHNSVRALIPAPPLPSQTDGAGARNQLINSNGMSRLFTPRNSVADAA 60
QY 60 DNRAADVGLPVNPRMLAASEITLNDGFEVLHDHGPLDTLNRLQISSVFRVETORDGKH 119
DB 61 DSRADTGLPTNPLRPAASEVSLHGALFVLDHDKGLDTLNSAISLFRVETRDGSHV 120
QY 120 AVGQNGVETSVVLSDDYARLQSIDPEKDKFVTTGGGAGGAMVTVASDITTEARQRI 179
DB 121 AIGQRNGLETTVLSSEOFSSLSQSLDPEKKNKFVTTGGGAGGAMVTVASDIARQRI 180
QY 180 LELPKGTGSEKSGAGESKGVGELRESNGAENTTETOTSTSSLRSDPKMLALGTVA 239
DB 181 IDKLEPKOTKETKEG-----DPSGEGKIIEHTSTSTSSLRADPKMLSLGITA 231
QY 240 TGLIGLAATGIVQALALTPEDPSPTTTPDDAASAETATRDQLTKEAFQNDQKVNID 299
DB 232 AGLIGMAATGIAQVAALTPEDPDPTTDDAANATAAAXKQLTKEARQNDQKVNID 291
QY 300 EIGNAIPESGVLKDDVYANIEBQAKAAGEAKQQAIEENNAQAKTDEQAKRQBELKYS 359
DB 292 EGNNAIPSELKDDVVAQIAEQAKAAGEARQEAIESNSQAQOKDEQAKRQBELKYS 351
QY 360 GAGYGISGALIIGGGIGVAVTAALHRKQPVETQTTTTTTTTTSARTYENKPAANTPQ 419
DB 352 GVGYGISGALIIGGGIGVAVTAALHRKQPAEQTTTTT-----VVDNPTNNAQA 403
QY 420 GNVDTPGSEDTMESRRSSVASTSTPTFDTSIGTVQNPYADYKSLHDSQVETSNSTSV 479
DB 404 GNTDTSGPESPASRRNSVASLASNSGSDTSSGTVENPADV-----GMRNDSIARI 456
QY 480 QNMGNTDSVY---YSTIOPPRDITDNGARLLGNPSAGIYSTYARLALSGLRHDMGSL 535
DB 457 SEEPYDEVAADPNYSVIOHFGSGNSPVTG-RIVGTGQGIQSTYVALLASGGRLRMGGL 515

```

```

QY 536 TCGSNSAVNTSNPPAPGSHRFV 558
DB 516 TGGGESAVNSTANAPTPGPAPFV 538

RESULT 8
ID O5K5P9 ECOLI PRELIMINARY; PRT; 538 AA.
AC O5K5P9
DT 15-FEB-2005, integrated into UniProtKB/TrEMBL.
DT 15-FEB-2005, sequence version 1.
DT 07-FEB-2006, entry version 4.
DE Translocated intimin receptor.
GN Name=tir;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=RM1374;
RX MEDLINE=21536660; PubMed=11682182;
RA Jores J., Rumer L., Kieseiling S., Kaper J.B., Wieler L.H.;
RT "Identification of a new pathogenicity island inserted in the phev
RT tRNA gene of the bovine Shiga toxin-producing E. coli strain RM1374
RT (O103:H2) harboring a locus of enterocyte effacement that is flanked
RT by intact insertion elements";
RL FEMS Microbiol. Lett. 204:75-79(2001).
RN
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=RM1374;
RX MEDLINE=22522600; PubMed=12635929;
RA Rumer L., Jores J., Kirsch P., Cavignac Y., Zehmke K., Wieler L.H.;
RT "Dissemination of phev and phev located genomic islands among
RT enteropathogenic (EPEC) and enterohemorrhagic (EHEC) E. coli and their
RT possible role in the horizontal transfer of the locus of enterocyte
RT effacement (LEE).";
RL Int. J. Med. Microbiol. 292:463-475(2003).
RN
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=RM1374;
RA Jores J., Wagner S.K., Rumer L., Eichberg J., Latynus C., Kirsch P.,
RA Schierack P., Tschape H., Wieler L.H.;
RT "Description of a 111-kb pathogenicity island (PAI) encoding various
RT virulence features in the enterohemorrhagic E. coli (EHEC) strain
RT RM1374 (O103:H2) and detection of a similar PAI in other EHEC strains
RT of serotype O103:H2.";
RL Int. J. Med. Microbiol. 294:417-425(2005).
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs license
CC
CC -----
CC EMBL; AJ303141; CA143867.1; -; Genomic_DNA.
CC SMR; O5K5P9; 261-325.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR003536; T1r rcpt.
DR Pfam; PF07489; T1r_receptor_C; 1.
DR Pfam; PF03549; T1r_receptor_M; 1.
DR Pfam; PF07490; T1r_receptor_N; 1.
DR PRINTS; PR01370; TRANSINTIMINR.
KM Receptor.
SQ SEQUENCE 538 AA; 55482 MW; 7F05DB3C6207F117 CRC64;

Query Match 64.6%; Score 1835; DB 2; Length 538;
Best Local Similarity 66.1%; Pred. No. 3.9e-89;
Matches 371; Conservative 60; Mismatches 104; Indels 26; Gaps 6;

QY 1 MPIGLGNPNVNNISIPAPPLPSQTDGAGG-RGQLINSTGPGLSRALFTPRNSMADSG 59
DB 1 MPIGLGNHNSVRALIPAPPLPSQTDGAGARNQLINSNGMSRLFTPRNSVADAA 60
QY 60 DNRAADVGLPVNPRMLAASEITLNDGFEVLHDHGPLDTLNRLQISSVFRVETORDGKH 119

```

```

Db      61 DSRASDIPGLPTNPRFAASEVSLHGALEVLHDKGLDPTLNSAIGSSLFRVETRDGSHV 120
Oy      120 AVGORNQVETSVVLSDOEYARLQSIDPEKDKFVFTEGCGAGHAMVTVASDITTEARQRI 179
Db      121 AIGQKNGLETTVVLSDQERSSLSQSIDPEKKNKFVFTGCGAGHAMVTVASDITTEARQRI 180
Oy      180 LELLEPKGTGSKGAGESKGVGELREBSNGAENTTETOTSTSSLSRSDPKMLALGTVA 239
Db      181 IDKLEPKDKETKEPG-----DPSNGEGKIIIEIHTSTSSLSRADPKMLALGTVA 231
Oy      240 TGLIGLATGIVQALALTEPDSPTTTPDPAASATETATRDQITKEAFQNDPNQKVID 299
Db      232 AGLIEMAATGIAQVAALTEPDDPTTTPDPAASATETATRDQITKEAFQNDPNQKVID 291
Oy      300 ELGNAPSGVLKDDVAVANIEBOAKAAGEAKQAQIENNAQAOKKYDEQAKQOEELKVS 359
Db      292 ENGNAIPSGELIDVVAQIAEQAKAAGEARQEALESNSQAOKKYDEQAKQOEELKVS 351
Oy      360 GAGYGLSGALLIGGIGVAVTAALHRKNQPVETQTTTTTTTTTSARTVENKPAANTPAQ 419
Db      352 GVGYIGSGLIIGGIGAGVTAALHRKNQPAEQITITRT-----VVDNQPTNNAQAQ 403
Oy      420 GNVDPGSEDIMESRRSSMASTSTFPDTSIGTVQNPYADVKYSLHDS--OVPTSNSMT 477
Db      404 GNTDTSGEESPASRRSNASLASNGSDTSTGYENPVADVGMPRNDLSARIPEPIYD 463
Oy      478 SVQNMGNNTDSVYVSTIQHPPTDNGARLGNPSAGIQTAYARLASGGLRHMGGLTG 537
Db      464 EVAADPN-----YSYIOHPSGNSPVYTG-RLVGTPEGQIGQISTYALLASGGLRLMGGLTG 517
Oy      538 GNSAIVNTSNPPAPGSHRFV 558
Db      518 GGSASVSTANASPTPGPARFV 538

RESULT 9
O58189_ECOLI PRELIMINARY; PRT; 538 AA.
AC O58189;
DT 26-APR-2005, integrated into UniProtKB/TrEMBL.
DT 26-APR-2005, sequence version 1.
DT 07-FEB-2006, entry version 6.
DE Translocated intimin receptor.
GN Name=tir;
OS Escherichia coli O157:H-.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=183192;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Gwynns Run Gwynnbrook stream;
RX PubMed=15870341; DOI=10.1128/AEM.71.5.2511-2519.2005;
RA Higgins J.A., Belt K.T., Karns J.S., Russell-Anelli J., Shelton D.R.;
RT "tir- and stx-Positive Escherichia coli in Stream Waters in a
  Metropolitan Area.";
RL Appl. Environ. Microbiol. 71:2511-2519(2005).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Gwynns Run Gwynnbrook stream;
RA Hohn C., Shelton D.R., Higgins J.A.;
RL Submitted (FEB-2005) to the EMBL/GenBank/DBJ databases.

-----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs license
CC
EMBL: AY944736; AAX47729.1; -; Genomic_DNA.
SMR: O58189; 261-325.
DR GO: GO:0004872; F:receptor activity; IEA.
DR GO: GO:0007155; P:cell adhesion; IEA.
DR InterPro: IPR003536; TIR.rcpt.
DR Pfam: PF07489; TIR_receptor_C; 1.
DR Pfam: PF03549; TIR_receptor_M; 1.

```

```

DR Pfam: PF07490; TIR_receptor_N; 1.
DR PRINTS: PR01370; TRNSINTIMINR.
KW Receptor.
SQ SEQUENCE 538 AA; 55496 MW; 9B848C2F508FC943 CRC64;
Query Match 64.5%; Score 1832; DB 2; Length 538;
Best Local Similarity 66.1%; Pred. No. 5,6e-89;
Matches 371; Conservative 60; Mismatches 104; Indels 26; Gaps 6;
Oy      1 MPTGNLGNPNVNNISPPAPPLPSQTDGAG--RGQLINSTGPGSALTPPVNSNADSG 59
Db      1 MPTGNLGNPNVNNISPPAPPLPSQTDGAG--RGQLINSTGPGSALTPPVNSNADSG 60
Oy      60 DRRASDVPELPVNPBRLAASEITLNDGEFVLHDHGLDPTLNSAIGSSLFRVETRDGSHV 119
Db      61 DSRASDIPGLPTNPRFAASEVSLHGALEVLHDKGLDPTLNSAIGSSLFRVETRDGSHV 120
Oy      120 AVGORNQVETSVVLSDOEYARLQSIDPEKDKFVFTEGCGAGHAMVTVASDITTEARQRI 179
Db      121 AIGQKNGLETTVVLSDQERSSLSQSIDPEKKNKFVFTGCGAGHAMVTVASDITTEARQRI 180
Oy      180 LELLEPKGTGSKGAGESKGVGELREBSNGAENTTETOTSTSSLSRSDPKMLALGTVA 239
Db      181 IDKLEPKDKETKEPG-----DPSNGEGKIIIEIHTSTSSLSRADPKMLALGTVA 231
Oy      240 TGLIGLATGIVQALALTEPDSPTTTPDPAASATETATRDQITKEAFQNDPNQKVID 299
Db      232 AGLIEMAATGIAQVAALTEPDDPTTTPDPAASATETATRDQITKEAFQNDPNQKVID 291
Oy      300 ELGNAPSGVLKDDVAVANIEBOAKAAGEAKQAQIENNAQAOKKYDEQAKQOEELKVS 359
Db      292 ENGNAIPSGELIDVVAQIAEQAKAAGEARQEALESNSQAOKKYDEQAKQOEELKVS 351
Oy      360 GAGYGLSGALLIGGIGVAVTAALHRKNQPVETQTTTTTTTTTSARTVENKPAANTPAQ 419
Db      352 GVGYIGSGLIIGGIGAGVTAALHRKNQPAEQITITRT-----VVDNQPTNNAQAQ 403
Oy      420 GNVDPGSEDIMESRRSSMASTSTFPDTSIGTVQNPYADVKYSLHDS--OVPTSNSMT 477
Db      404 GNTDTSGEESPASRRSNASLASNGSDTSTGYENPVADVGMPRNDLSARIPEPIYD 463
Oy      478 SVQNMGNNTDSVYVSTIQHPPTDNGARLGNPSAGIQTAYARLASGGLRHMGGLTG 537
Db      464 EVAADPN-----YSYIOHPSGNSPVYTG-RLVGTPEGQIGQISTYALLASGGLRLMGGLTG 517
Oy      538 GNSAIVNTSNPPAPGSHRFV 558
Db      518 GGSASVSTANASPTPGPARFV 538

RESULT 10
O85508_ECOLI PRELIMINARY; PRT; 538 AA.
AC O85508;
DT 01-NOV-1998, integrated into UniProtKB/TrEMBL.
DT 01-NOV-1998, sequence version 1.
DT 07-FEB-2006, entry version 16.
DE Translocated intimin receptor Tir.
GN Name=tir;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=EPEC87A;
RX MEDLINE=99003184; PubMed=9784578;
RA Paton A.W., Manning P.A., Woodrow M.C., Paton J.C.;
RT "Translocated intimin receptors (tir) of Shiga-toxinogenic Escherichia
  coli isolates belonging to serogroups O26, O111, and O157 react with
  sera from patients with hemolytic-uremic syndrome and exhibit marked
  sequence heterogeneity.";
RL Infect. Immun. 66:5580-5586(1998).

```

CC -----
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC distributed under the Creative Commons Attribution-NonDerivs License
CC
EMBL; AF070069; AAC69318.1; -; Genomic_DNA.
DR HSSP; Q9KWH9; 1f02.
DR SMR; O85508; 261-325.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR003536; TIR_rept.
DR Pfam; PF07489; TIR_receptor_C; 1.
DR Pfam; PF03549; TIR_receptor_M; 1.
DR Pfam; PF07490; TIR_receptor_N; 1.
DR PRINTS; PR01370; TRANSINTIMINR.
KW Receptor.
SQ SEQUENCE 538 AA; 55602 MW; 447052A0E3214D6D CRC64;

Query Match 64.2%; Score 1824; DB 2; Length 538;
Best Local Similarity 65.8%; Pred. No. 1.5e-88;
Matches 369; Conservative 60; Mismatches 106; Indels 26; Gaps 6;

QY 1 MPINLGNPNVNSIPAPPLPSQTGAGC-RGLINSTGPLSRALFTPRNSMADSG 59
DB 1 MPINLGNPNVNRALIPAPPLPSQTGAGARNQLINSNGMGRLLFTPIRNSVADAA 60
QY 60 DNRAADVGLPVNPMRLAASBITLNDGFEVLHDGPLDTLNRQISSVFRVETOEDGKH 119
DB 61 DSRASDIDGLPTNPLRFAASEVSLHGALEVLHDGGLDTLSALIGSSLFRVETRDGSHV 120
QY 120 AVGORNVEYSVLSDOEYARLQSIDPEGKDFVFTGSGAGHAMVTVAASDITEARQRI 179
DB 121 AIGQNGLETTVVLSDQEFSSIQSLDPEGKKNFVFTGSGAGHAMVTVAASDIAEARQRI 180
QY 180 LELLEPKTGSKGAGESKVGELRESNGAENTTETOTSTSTSLRSDPKMLALGTV 239
DB 181 IDKLEPKDTKETKEG-----DPNSGEGKIIIEHTSTSTSLRADPKMLSLGTIA 231
QY 240 TGLGLAATGIVQALALTPPEPDSPTTTPDPAASAETATRDOLTKFAFQNPDNOKNID 299
DB 232 AGLIGMAATGIAQVALTPPEPDDPTTTPDPAASAEATKRLQGEAFQDPDKKNID 291
QY 300 ELGNAIPSGVLKDVVANIIEQAKAAGEAKQAQIENNAOQKYDEQAKROBELKYSS 359
DB 292 ENGNAIPSGELIDVVAQIAEQAKAAGEAQGEALIESNSQAKKYDEQAKREGEMLSS 351
QY 360 GAGYGLSGLILGGIGIAVVAALAHKKNQPEQTTTTTTTTTSARVYENKPNANTPAQ 419
DB 352 GVGYGISGAILVGGIGIAGVVAALAHKKNQPEQTTTTT-----VVDNQPTNNSAQ 403
QY 420 GNVDPGSEDTMESRSSMASTSTFPPTSSIGTVQNPYADVKTSLHDS--OVPTSNT 477
DB 404 GNTDTSGEESPAARRNSASLASNGSTSTGIVENPYADVGMRRNSLARIPEEPIYD 463
QY 478 SVQNGNTDSVVYSTIQHPPRDITDNGARLLGNPSAGIQSTYARLALSGGLRHDGGL 537
DB 464 EVAADPN-----YSTIQHPSGNSPVLTG-RLVGTPEGGIQSTYALLASSGGLRLGNGGL 517
QY 538 GSNSAVNTSNPPAPGSHRFV 558
DB 518 GGESAVSTANASPTPGPARFV 538

RESULT 11
058187_ECOLI PRELIMINARY; PRT; 538 AA.
AC 058187;
DT 26-APR-2005, integrated into UniProtKB/TrEMBL.
DT 26-APR-2005, sequence version 1.
DE 07-FEB-2006, entry version 6.
DE Translocated intimin receptor.
GN Name=Iir;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15870341; DOI=10.1128/AEM.71.5.2511-2519.2005;
RA Higgins J.A., Belt K.T., Kame J.S., Russell-Anelli J., Shelton D.R.;
RT "Iir" and sex-Positive Escherichia coli in Stream Waters in a
Metropolitan Area."
RU Appl. Environ. Microbiol. 71:2511-2519(2005).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Holm C., Kame J.S., Higgins J.A.;
RA Submitted (FEB-2005) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC distributed under the Creative Commons Attribution-NonDerivs License
CC
EMBL; AY944738; AAK47731.1; -; Genomic_DNA.
DR SMR; O58187; 261-325.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR003536; TIR_rept.
DR Pfam; PF07489; TIR_receptor_C; 1.
DR Pfam; PF03549; TIR_receptor_M; 1.
DR Pfam; PF07490; TIR_receptor_N; 1.
DR PRINTS; PR01370; TRANSINTIMINR.
KW Receptor.
SQ SEQUENCE 538 AA; 55455 MW; D28F5200F04A1890 CRC64;

Query Match 64.1%; Score 1820; DB 2; Length 538;
Best Local Similarity 65.4%; Pred. No. 2.4e-88;
Matches 368; Conservative 62; Mismatches 103; Indels 30; Gaps 6;

QY 1 MPINLGNPNVNSIPAPPLPSQTGAGC-RGLINSTGPLSRALFTPRNSMADSG 59
DB 1 MPINLGNPNVNRALIPAPPLPSQTGAGARNQLINSNGMGRLLFTPIRNSVADAA 60
QY 60 DNRAADVGLPVNPMRLAASBITLNDGFEVLHDGPLDTLNRQISSVFRVETOEDGKH 119
DB 61 DSRASDIDGLPTNPLRFAASEVSLHGALEVLHDGGLDTLSALIGSSLFRVETRDGSHV 120
QY 120 AVGORNVEYSVLSDOEYARLQSIDPEGKDFVFTGSGAGHAMVTVAASDITEARQRI 179
DB 121 AIGQNGLETTVVLSDQEFSSIQSLDPEGKKNFVFTGSGAGHAMVTVAASDIAEARQRI 180
QY 180 LELLEPKTGSKGAGESKVGELRESNGAENTTETOTSTSTSLRSDPKMLALGTV 239
DB 181 IDKLEPKDTKETKEG-----DPNSGEGKIIIEHTSTSTSLRADPKMLSLGTIA 231
QY 240 TGLGLAATGIVQALALTPPEPDSPTTTPDPAASAETATRDOLTKFAFQNPDNOKNID 299
DB 232 AGLIGMAATGIVQVALTPPEPDDPTTTPDPAASAEATKRLQGEAFQDPDKKNID 291
QY 300 ELGNAIPSGVLKDVVANIIEQAKAAGEAKQAQIENNAOQKYDEQAKROBELKYSS 359
DB 292 ENGNAIPSGELIDVVAQIAEQAKAAGEAQGEALIESNSQAKKYDEQAKREGEMLSS 351
QY 360 GAGYGLSGLILGGIGIAVVAALAHKKNQPEQTTTTTTTTTSARVYENKPNANTPAQ 419
DB 352 GVGYGISGAILVGGIGIAGVVAALAHKKNQPEQTTTTT-----VVDNQPTNNSAQ 403
QY 420 GNVDPGSEDTMESRSSMASTSTFPPTSSIGTVQNPYADVKTSLHDSOVPTSNSNTS 479
DB 404 GNTDTSGEESPAARRNSASLASNGSTSTGIVENPYADV-----GMRNDSLARI 456
QY 480 QNMGNTDSV-----YSTIQHPPRDITDNGARLLGNPSAGIQSTYARLALSGGLRHDGGL 535
DB 457 SEEPIYDEVADPNVSVIHFSGNSPVLTG-RLVGTPEGGIQSTYALLASSGGLRLGNGGL 515
QY 536 TGSNSAVNTSNPPAPGSHRFV 558
DB 516 TGGESAVSTANAAPTPGPARFV 538

```

RESULT 12
058190 ECOLI PRELIMINARY; PRT; 538 AA.
AC 058190;
DT 26-APR-2005, integrated into UniProtKB/TrEMBL.
DT 26-APR-2005, sequence version 1.
DT 07-FEB-2006, entry version 6.
DE Translocated intimin receptor.
GN Name: tir;
OS Escherichia coli O157:H-.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxId=183192;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Spring Branch stream;
RX PubMed=15870341; DOI=10.1128/AEM.71.5.2511-2519.2005;
RA Higgins J.A., Belt K.T., Karns J.S., Russell-Anelli J., Shelton D.R.,
RT "tir- and stx-Positive Escherichia coli in Stream Waters in a
RT Metropolitan Area.";
RL Appl. Environ. Microbiol. 71:2511-2519(2005).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Spring Branch stream;
RA Hohn C., Shelton D.R., Higgins J.A.;
RL Submitted (FEB-2005) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs license
CC -----
EMBL AY944735, AX47728.1; -, Genomic_DNA.
DR SMR; 058190; 261-325.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR003536; TIR rcpt.
DR Pfam; PF07489; TIR_receptor_C; 1.
DR Pfam; PF03549; TIR_receptor_M; 1.
DR Pfam; PF07490; TIR_receptor_N; 1.
DR PRINTS; PRO1370; TRANSINTIMINR.
DR Receptor.
SQ SEQUENCE 538 AA; 55609 MW; 0336B5E18787C18E CRC64;

Query Match 64.1%; Score 1820; DB 2; Length 538;
Best Local Similarity 66.0%; Pred. No. 2.4e-88;
Matches 370; Conservative 60; Mismatches 105; Indels 26; Gaps 6;

QY 1 MPIGNLGNPNVNNNSIPPAPPLPSQTDGAGG-RGOLINSTGPLGSRALFTPRNSMADSG 59
DB 1 MPIGNLGNPNVNRALIPPAPPLPSQTDGAGARNOQLINSGPWGRLFTPIRNSVADNA 60
QY 60 DNRAADVPGLPVNPMLAASEITLNDGFEVLHDGFLDPLTNROIGSSVFRVETQEDGKH 119
DB 61 DSRASDIPGLPTNPLRFAASEVSLHGALEVLHDKGLDPLTNLSIGSLRFRVETRDGSHV 120
QY 120 AAGQNGVTSVVLNQEYARLQSIDPEKQKXFFVFGGAGGAGHAMVTYASDITTEARQRI 179
DB 121 AIGQNGLETTVLVLNQEFSSLSQSLDPEKKNKFFVFGGAGGAGHAMVTYASIAEARQRI 180
QY 180 LELLPEKGTGSGKAGESKGVGELRESNGAENTTEOTSTSSLSRDPKMLALGTA 239
DB 181 IDKLEPKOTKETKEG-----DPSGEGKIIEHTSTSTSSLRAPKMLSLGITA 231
QY 240 TGLIGLATGIVQALALTPEDPSPTTDPDAASATETATRTDQLTKEAFQNDQKVNID 299
DB 232 AGLIGMAATGIVQAVALTPEPDDLTFTDPTAASTEAATKXRLTQEARQDPDKQVND 291
QY 300 ELGNAPISGVLDKDVANIEEQAAGAAEAKQAQIENNAQAQKTDDEQAQKEELKYS 359
DB 232 ELGNAPISSELIDDDVAQIAEQAKAGAEQARQEALESNSQAQKTDDEQAQKEEALSS 351
QY 360 GAGYGLSGALLGGIGVAVTAALHRKNQPVETQTTTTTTTTTTSARTVENKPAANTPAQ 419

```

```

DB 352 GVCYRISGALLGGIGAGVTAALHRKNQPAEQITTRT-----VVDNQPTNNAQA 403
QY 420 GNVDTFGSEDTMESRSSMASTSTFPDTSICTVONPIADVTSLHDS--QVTSNSNT 477
DB 404 GNVDTSGPEESPASRRNSASLASNGSDTSSTGTVENPVDVGMPPNDISARIPEEPIYD 463
QY 478 SYVNMGNTDSVYVSTIQHPROTDNGARLGNPSAGISQTVARLALSGLRHDMGLTG 537
DB 464 EVAADPN-----YSVQHPFSGNSPVTG-RLVGTPEGQIGISTYALLASGGLRLMGGLTG 517
QY 538 GSNVAVNTSNPPAPGSHRFV 558
DB 518 GGSNAVSTNASTPTPPAPFV 538

RESULT 13
068258 ECOLI PRELIMINARY; PRT; 551 AA.
AC 068258;
DT 01-AUG-1998, integrated into UniProtKB/TrEMBL.
DT 01-AUG-1998, sequence version 1.
DT 07-FEB-2006, entry version 19.
DE Translocated intimin receptor.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxId=562;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=98187918; PubMed=9529069;
RA Voss E., Paton A.W., Manning P.A., Paton J.C.;
RT "Molecular analysis of Shiga toxinogenic Escherichia coli O111:H-
RT proteins which react with sera from patients with hemolytic-uremic
RT syndrome.";
RL Infect. Immun. 66:1467-1472(1998).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=99003184; PubMed=9784578;
RA Paton A.W., Manning P.A., Woodrow M.C., Paton J.C.;
RT "Translocated intimin receptor (Tir) of Shiga-toxinogenic Escherichia
RT coli isolates belonging to serogroups O26, O111, and O157 react with
RT sera from patients with hemolytic-uremic syndrome and exhibit marked
RT sequence heterogeneity.";
RL Infect. Immun. 66:5580-5586(1998).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs license
CC -----
EMBL AF025311; AAC69249.1; -, Genomic_DNA.
DR HSSP; G9KWH9; 1F02.
DR SMR; 068258; 272-336.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR003536; TIR rcpt.
DR Pfam; PF07489; TIR_receptor_C; 1.
DR Pfam; PF03549; TIR_receptor_M; 1.
DR Pfam; PF07490; TIR_receptor_N; 1.
DR PRINTS; PRO1370; TRANSINTIMINR.
DR Receptor.
SQ SEQUENCE 551 AA; 56975 MW; 6EC95F76E0F44CC CRC64;

Query Match 56.8%; Score 1613.5; DB 2; Length 551;
Best Local Similarity 60.8%; Pred. No. 2.2e-77;
Matches 346; Conservative 68; Mismatches 126; Indels 29; Gaps 13;

QY 1 MPIGNLGNPNVNNNSIPPAPPLPSQTDGA--GGRQLNSTGPLGSRALFTPRNSMADS 58
DB 1 MPIGNLGNPNVNNNLIPPAPPLPSQTDASRGAAQLINSGALSLRFLFSPLRNSIDS 60
QY 59 GDNRAADVPGLPVNPMLR--AASEITLNDGFEVLHDGFLDPLTNROIGSSVFRVETQEDG 116
DB 61 VDSR--DIPGLPVNHRSLATATSEICLGGFEVLHDKGPDLTNLQIGASAFRIQGSNG 118

```

```

QY 117 KHAIVGQRNGEYVSVLSIDOEYARLOSIDPEGKDFVFTGGGAGAHAMVTASDITEAR 176
D 119 SYAIAEGKNGEYVSVILNSQELQSLQALDIEDKGFVFTGGGAGAHAMVTASDITEAR 178
QY 177 ORIELLEPKGTGSKAG--ESKVGELRESNSGANT--TETQSTSTSSLSRDPKL 231
D 179 AKIIAKLDPNNHGSGQARNVDTRSVGVG---SASGMDSVVSETRTSSTASVRSDEKF 234
QY 232 WLAGTVATGLIGLAATGIVQALATPEPDSPTTTPDPAASATETATRDOLTEAFQNP 291
D 235 WVSIGALAGLAGLAATGITQALATPEPDDPTTTPDPAASABSATRDOLTEAFQNP 294
QY 292 DNQKVNIDELGNALPSGVLKQDVVANIIEQAKAGEAKQQAIEENNAQOKKYDEQAKR 351
D 295 ENQKVSIDEIGNSIPISEGLKQDVVAKIEQAKAGEAKQQAIVESNAQOQRYPDYARR 354
QY 352 QEBLKVSAGAGYGLSGLILGCGIGAVYTAALHRKNQVEQTTTTTTTTTTTSAATYENK 411
D 355 QEBLELSGIGISLSALIVGGIGAGVYTTALHRNQPABEQTTTTTHT-----VVOQQ 408
QY 412 PANTPAGQNDTPGSEDTMSESRSSMASTSTPF--DTSSICTYONPYADYKTSIHDQV 470
D 409 TGGNTPAQGTDALRAEDTSLNRDSQSTASTHSDTSS--AVNPPAIEGEARNSS-- 464
QY 471 PTSNSNTSVQNMGNMTDSVYVSTIQHPPRDTTNGARLLGNPSAGIOSTYARLA--LSGLR 529
D 465 PARQAEHIIDEVADP--NYSVIONFGSNNQVTG--RLMGTGQGGIGQSTYALITNNSAGLR 522
QY 530 HDMGGLTGSNSAVNTSNPPAPGSHRFV 558
D 523 LGMGGLTGSNGSAAVNTANAAPTPGPRFV 551

RESULT 14
Q42IM1_ECOLI PRELIMINARY; PRT; 551 AA.
AC Q42IM1;
DT 07-JUN-2005, integrated into UniProtKB/TrEMBL.
DT 07-JUN-2005, sequence version 1.
DE 07-FEB-2006, entry version 4.
DE Translocated intimin receptor.
GN Name=tir;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OC NCBI_TaxID=562;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CPG121;
RX PubMed16272509; DOI=10.1128/JCM.43.11.5715-5720.2005;
RA Gammedia J., Ren Z., Tennant S., Midolli Viera M.A., Chong Y.,
RA Whale A., Azopardi K., Dahan S., Sircili M.P., Franzolin M.R.,
RA Trabulsi L.R., Phillips A., Gomes T.A., Xu J., Robins-Browne R.,
RA Frankel G.;
RT "Distribution of tccp in Clinical Enterohemorrhagic and
RT Enteropathogenic Escherichia coli Isolates.";
RL J. Clin. Microbiol. 43:5715-5720(2005).
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
DR EMBL: DQ007023; AAY2394.1; -; Genomic_DNA.
DR SMR: Q42IM1: 272-336.
DR GO: GO:0004872; F:receptor activity; IEA.
DR GO: GO:0007155; P:cell adhesion; IEA.
DR InterPro: IPR003536; TIR:rcpt.
DR Pfam: PF07489; TIR_receptor_C; 1.
DR Pfam: PF07490; TIR_receptor_M; 1.
DR PRINTS: PR01370; TRANSITIMIR.
DR Receptor.
SQ SEQUENCE 551 AA; 57066 MW; 333DB01FC9B461CB CRC64;

```

```

Query Match 56.7%; Score 1611.5; DB 2; Length 551;
Best Local Similarity 60.9%; Pred. No. 2,86-77;
Matches 346; Conservative 67; Mismatches 126; Indels 29; Gaps 13;

QY 1 MPINLGNPNVNSIIPAPLPBQTDCA--GGGQLINSGPLGSRALFPVNRSMADS 58
D 1 MPINLGNPNVNSNLIIPAPLPBQTDGASNGAGOLINSGALGSRLLFPPLNRSIADS 60
QY 59 GDNBASDPVGPVPMRL--AASEITLNDGEVYHDGPLTLNRQIGSSVFRVETQEDG 116
D 61 VDSR--DIPGLPVPSRLATTSLSCLLGEFEVYHDGPLTLNKQIGASAFRIEQSDG 118
QY 117 KHAIVGQRNGEYVSVLSIDOEYARLOSIDPEGKDFVFTGGGAGAHAMVTASDITEAR 176
D 119 SYAIAEGKNGEYVSVILNSQELQSLQALDIEDKGFVFTGGGAGAHAMVTASDITEAR 178
QY 177 ORIELLEPKGTGSKAG--ESKVGELRESNSGANT--TETQSTSTSSLSRDPKL 231
D 179 AKIIAKLDPNNHGSGSKADVDTRSVGVG---SASGMDSVVSETRTSSTASVRSDEKF 234
QY 232 WLAGTVATGLIGLAATGIVQALATPEPDSPTTTPDPAASATETATRDOLTEAFQNP 291
D 235 WVSIGALAGLAGLAATGITQALATPEPDDPTTTPDPAASABSATRDOLTEAFQNP 294
QY 292 DNQKVNIDELGNALPSGVLKQDVVANIIEQAKAGEAKQQAIEENNAQOKKYDEQAKR 351
D 295 ENQKVSIDEIGNSIPISEGLKQDVVAKIEQAKAGEAKQQAIVESNAQOQRYPDYARR 354
QY 352 QEBLKVSAGAGYGLSGLILGCGIGAVYTAALHRKNQVEQTTTTTTTTTTTSAATYENK 411
D 355 QEBLELSGIGISLSALIVGGIGAGVYTTALHRNQPABEQTTTTTHT-----VVOQQ 408
QY 412 PANTPAGQNDTPGSEDTMSESRSSMASTSTPF--DTSSICTYONPYADYKTSIHDQV 470
D 409 TGGNTPAQGTDALRAEDTSLNRDSQSTASTHSDTSS--AVNPPAIEGEARNSS-- 464
QY 471 PTSNSNTSVQNMGNMTDSVYVSTIQHPPRDTTNGARLLGNPSAGIOSTYARLA--LSGLR 529
D 465 PARQAEHIIDEVADP--NYSVIONFGSNNQVTG--RLMGTGQGGIGQSTYALITNNSAGLR 522
QY 530 HDMGGLTGSNSAVNTSNPPAPGSHRF 557
D 523 LGMGGLTGSNGSAAVNTANAAPTPGPVRF 550

RESULT 15
Q42IM0_ECOLI PRELIMINARY; PRT; 551 AA.
AC Q42IM0;
DT 07-JUN-2005, integrated into UniProtKB/TrEMBL.
DT 07-JUN-2005, sequence version 1.
DE 07-FEB-2006, entry version 5.
DE Translocated intimin receptor.
GN Name=tir;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OC NCBI_TaxID=562;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CPG124-MC416, and CPG123-G58;
RX PubMed16272509; DOI=10.1128/JCM.43.11.5715-5720.2005;
RA Gammedia J., Ren Z., Tennant S., Midolli Viera M.A., Chong Y.,
RA Whale A., Azopardi K., Dahan S., Sircili M.P., Franzolin M.R.,
RA Trabulsi L.R., Phillips A., Gomes T.A., Xu J., Robins-Browne R.,
RA Frankel G.;
RT "Distribution of tccp in Clinical Enterohemorrhagic and
RT Enteropathogenic Escherichia coli Isolates.";
RL J. Clin. Microbiol. 43:5715-5720(2005).
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----

```

DR EMBL; DQ007024; AAY25395.1; -; Genomic DNA.
DR EMBL; DQ007022; AAY25393.1; -; Genomic DNA.
DR SMR; Q421M0; 272-336.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0007155; F:cell adhesion; IEA.
DR InterPro; IPR003536; T1r_rcpl.
DR Pfam; PF07489; T1r_receptor_C; 1.
DR Pfam; PF03549; T1r_receptor_M; 1.
DR Pfam; PF07490; T1r_receptor_N; 1.
DR PRINTS; PR01370; TENSINTIMINR.
KW Receptor.
SQ SEQUENCE 551 AA; 57066 MW; 83B899E69183AFB4 CRC64;

Query Match 56.6%; Score 1607.5; DB 2; Length 551;
Best Local Similarity 60.7%; Pred. No. 4,6e-77;
Matches 345; Conservative 68; Mismatches 126; Indels 29; Gaps 13;

QY 1 MPIGNIGNHPNVNNSIPAPPLPSQTDG--GGRGQLINSTGPLGSRALFTPVNRMADS 58
DB 1 MPIGNIGNHPNVNNSIPAPPLPSQTDGASRGAGQLINSTGALGSRLLFSPLRNSIADS 60
QY 59 GGNRASDVDPGLPVPNPKRL--AASETLNDGFEVLHDHGPLDTLNKROIGSSVFRETQEDG 116
DB 61 VDSR--DIPGLPVHPSRLATATSEICLGFEVLHDHKGPLDTLNKROIGASAFRIEQQSDG 118
QY 117 KHIAVGQRNGVEYSVYLSDOEYARLOSIDPEGKDFVFTGGRGAGHAMWTVASDITTEAR 176
DB 119 STAAIEKNGVEYSVYLSDOEYARLOSIDPEGKDFVFTGGRGAGHAMWTVASDITTEAR 178
QY 177 QRILELLEPKGTGESKAG--ESKGVGEIRESNGAENT--TETOTSTSTSLRSDPKL 231
DB 179 AKILAKLDNNHGGSGARVDTRSVG---SASGMDSVSVSETRTSSASTASVRSDPKF 234
QY 232 WLALGTVAATGLGLAATGIVQALALTPPEPDSPTTTDPDAASATETATRDQLTKEAFQNP 291
DB 235 WVSIGAIAGLAGLAATGITOALALTPPEPDDPTTDPQOASASAFSATRDQLTQEAFAKNP 294
QY 292 DNOKVNIIDELGNAIPSGVLKDDVVAIEEQAKAAGEAKQOALENNAQOAKYDEQOAKR 351
DB 295 ENOKVSIDEIGNSIPSGELKDDVVAIEEQAKEAGEAKQOAVESNAQOQORYDTQYARR 354
QY 352 QBELKVSAGAGYGLSGALLGGGIGVAVTAAALHRKNQPVQPTTTTTTTTTTSARTVENK 411
DB 355 QBELKVSAGAGYGLSGALLGGGIGVAVTAAALHRKNQPVQPTTTTTTTTTTSARTVENK 411
QY 412 PANNTPAQGNVDTPGSEDTMESRRSSMASTSTFF-DTSSIGTVQNPYADVXTSLHDQV 470
DB 409 TCGNTPAQGTDAIRAEFTSLNRBDQSRSTASTHMDTSS--AVVPYAEVGEARRSS-- 464
QY 471 PTSNNTSVQNNNGNTDSVYVSTIQHPRTDTNGARLLGNPSAGIOSTYARLA-LSGGLR 529
DB 465 PARQAEHHYDEVADP-NYSVIQNSGNNQVYG-RLMGTPOGIOSTYAILTNNSAGLR 522
QY 530 HDMGGLTGGSNSAVNTSNPPAPGSHRF 557
DB 523 LGMGGLTGGSNSAVNTSNPPAPGSHRF 550

Search completed: August 1, 2006, 22:13:15
Job time : 307 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 1, 2006, 22:29:16 ; Search time 189 Seconds
(without alignments)
1349.877 Million cell updates/sec

Title: US-09-189-415D-11

Perfect score: 558

Sequence: 1 MPIGNIGHNPVNNPISPPAP.....SNSAVNTSNPPAPGSHRFV 558

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 2589679 seqs, 457216429 residues

Word size : 8

Total number of hits satisfying chosen parameters: 1176

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: listing first 1000 summaries

Database :

A_Geneseq.8:.*
1: geneseqp19808:.*
2: geneseqp19908:.*
3: geneseqp20008:.*
4: geneseqp20018:.*
5: geneseqp20028:.*
6: geneseqp20038:.*
7: geneseqp20038:.*
8: geneseqp20048:.*
9: geneseqp20058:.*
10: geneseqp20068:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	558	100.0	558	7	Adc00799 Enterohae
2	558	100.0	558	9	Aeb91310 Microbial
3	558	100.0	558	10	Aee86220 Escherich
4	452	81.0	559	2	Aay06221 EBEC E. C
5	14	2.5	30	2	Aay06213 EBEC E. C
6	14	2.5	549	2	Aay06220 EBEC E. C
7	13	2.3	162	10	Aee26469 ORF0826 i
8	13	2.3	166	6	Abj18927 Pathogen
9	13	2.3	166	6	Abj18927 Pathogen
10	13	2.3	166	10	Abm71151 Staphyloc
11	13	2.3	166	10	Aee26470 ORF0826 i
12	13	2.3	166	10	Aee26471 ORF0826 i
13	13	2.3	166	10	Aee26472 ORF0826 i
14	13	2.3	167	10	Aee26467 ORF0826 i
15	13	2.3	211	10	Aee26468 Full Leng
16	12	2.2	81	10	Aef20212 Pyrococcu
17	12	2.2	141	3	Aag10881 Arabidops
18	12	2.2	146	3	Aag18210 Arabidops
19	12	2.2	150	3	Aag10880 Arabidops
20	12	2.2	179	3	Aag10879 Arabidops
21	12	2.2	180	3	Aag18209 Arabidops
22	12	2.2	232	3	Aag18208 Arabidops
23	12	2.2	361	7	Adc86999 Human GPC
			548	9	Adv14340 Pyrococcu

24	12	2.2	561	4	ABE67878	Abb67878 Drosophi
25	12	2.2	785	8	ADP98986	Adp98986 C. albica
26	12	2.2	986	2	AAW59050	Aaw59050 H. contor
27	12	2.2	986	5	AAO17501	Aao17501 H. contor
28	12	2.2	986	5	AAO17502	Aao17502 H. contor
29	12	2.2	986	5	AAO17513	Aao17513 H. contor
30	12	2.2	1026	4	ABE64339	Abbe64339 Drosophi
31	12	2.2	1026	4	ABE64339	Abbe64339 Drosophi
32	12	2.2	1026	4	ABE64339	Abbe64339 Drosophi
33	12	2.2	1026	4	ABE64339	Abbe64339 Drosophi
34	12	2.2	1026	4	ABE64339	Abbe64339 Drosophi
35	12	2.2	1026	4	ABE64339	Abbe64339 Drosophi
36	12	2.2	1026	4	ABE64339	Abbe64339 Drosophi
37	12	2.2	1026	4	ABE64339	Abbe64339 Drosophi
38	12	2.2	1026	4	ABE64339	Abbe64339 Drosophi
39	12	2.2	1026	4	ABE64339	Abbe64339 Drosophi
40	12	2.2	1026	4	ABE64339	Abbe64339 Drosophi
41	12	2.2	1026	4	ABE64339	Abbe64339 Drosophi
42	12	2.2	1026	4	ABE64339	Abbe64339 Drosophi
43	12	2.2	1026	4	ABE64339	Abbe64339 Drosophi
44	12	2.2	1026	4	ABE64339	Abbe64339 Drosophi
45	12	2.2	1026	4	ABE64339	Abbe64339 Drosophi
46	12	2.2	1026	4	ABE64339	Abbe64339 Drosophi
47	12	2.2	1026	4	ABE64339	Abbe64339 Drosophi
48	12	2.2	1026	4	ABE64339	Abbe64339 Drosophi
49	12	2.2	1026	4	ABE64339	Abbe64339 Drosophi
50	12	2.2	1026	4	ABE64339	Abbe64339 Drosophi
51	12	2.2	1026	4	ABE64339	Abbe64339 Drosophi
52	12	2.2	1026	4	ABE64339	Abbe64339 Drosophi
53	12	2.2	1026	4	ABE64339	Abbe64339 Drosophi
54	12	2.2	1026	4	ABE64339	Abbe64339 Drosophi
55	12	2.2	1026	4	ABE64339	Abbe64339 Drosophi
56	12	2.2	1026	4	ABE64339	Abbe64339 Drosophi
57	12	2.2	1026	4	ABE64339	Abbe64339 Drosophi
58	12	2.2	1026	4	ABE64339	Abbe64339 Drosophi
59	12	2.2	1026	4	ABE64339	Abbe64339 Drosophi
60	12	2.2	1026	4	ABE64339	Abbe64339 Drosophi
61	12	2.2	1026	4	ABE64339	Abbe64339 Drosophi
62	12	2.2	1026	4	ABE64339	Abbe64339 Drosophi
63	12	2.2	1026	4	ABE64339	Abbe64339 Drosophi
64	12	2.2	1026	4	ABE64339	Abbe64339 Drosophi
65	12	2.2	1026	4	ABE64339	Abbe64339 Drosophi
66	12	2.2	1026	4	ABE64339	Abbe64339 Drosophi
67	12	2.2	1026	4	ABE64339	Abbe64339 Drosophi
68	12	2.2	1026	4	ABE64339	Abbe64339 Drosophi
69	12	2.2	1026	4	ABE64339	Abbe64339 Drosophi
70	12	2.2	1026	4	ABE64339	Abbe64339 Drosophi
71	12	2.2	1026	4	ABE64339	Abbe64339 Drosophi
72	12	2.2	1026	4	ABE64339	Abbe64339 Drosophi
73	12	2.2	1026	4	ABE64339	Abbe64339 Drosophi
74	12	2.2	1026	4	ABE64339	Abbe64339 Drosophi
75	12	2.2	1026	4	ABE64339	Abbe64339 Drosophi
76	12	2.2	1026	4	ABE64339	Abbe64339 Drosophi
77	12	2.2	1026	4	ABE64339	Abbe64339 Drosophi
78	12	2.2	1026	4	ABE64339	Abbe64339 Drosophi
79	12	2.2	1026	4	ABE64339	Abbe64339 Drosophi
80	12	2.2	1026	4	ABE64339	Abbe64339 Drosophi
81	12	2.2	1026	4	ABE64339	Abbe64339 Drosophi
82	12	2.2	1026	4	ABE64339	Abbe64339 Drosophi
83	12	2.2	1026	4	ABE64339	Abbe64339 Drosophi
84	12	2.2	1026	4	ABE64339	Abbe64339 Drosophi
85	12	2.2	1026	4	ABE64339	Abbe64339 Drosophi
86	12	2.2	1026	4	ABE64339	Abbe64339 Drosophi
87	12	2.2	1026	4	ABE64339	Abbe64339 Drosophi
88	12	2.2	1026	4	ABE64339	Abbe64339 Drosophi
89	12	2.2	1026	4	ABE64339	Abbe64339 Drosophi
90	12	2.2	1026	4	ABE64339	Abbe64339 Drosophi
91	12	2.2	1026	4	ABE64339	Abbe64339 Drosophi
92	12	2.2	1026	4	ABE64339	Abbe64339 Drosophi
93	12	2.2	1026	4	ABE64339	Abbe64339 Drosophi
94	12	2.2	1026	4	ABE64339	Abbe64339 Drosophi
95	12	2.2	1026	4	ABE64339	Abbe64339 Drosophi
96	12	2.2	1026	4	ABE64339	Abbe64339 Drosophi

97	11	2.0	216	3	AAB11729	Abb11729	Cryptospor	170	11	2.0	440	6	ABO02095	Abc02095	Human	sec
98	11	2.0	216	5	AB11734	Abb11734	Cryptospor	171	11	2.0	440	6	ABU58360	Abu58360	Novel	hum
99	11	2.0	216	5	ABJ04047	Abj04047	C parvum	172	11	2.0	440	6	ABU88669	Abu88669	Novel	hum
100	11	2.0	229	3	AB25593	Ab25593	Protein e	173	11	2.0	440	6	ABU83364	Abu83364	Human	sec
101	11	2.0	229	4	AB25593	Ab25593	Drosophill	174	11	2.0	440	6	ABO06165	AbO06165	Novel	hum
102	11	2.0	229	6	AB25593	Ab25593	Human nov	175	11	2.0	440	6	ABR59201	AbR59201	Human	sec
103	11	2.0	229	6	ADA27065	Ada27065	Human nov	176	11	2.0	440	6	ABO09263	AbO09263	Human	sec
104	11	2.0	229	8	AD25555	Ad25555	Novel hum	177	11	2.0	440	6	ABO19127	AbO19127	Novel	hum
105	11	2.0	239	4	ABG06102	Abg06102	Novel hum	178	11	2.0	440	6	ABO11145	AbO11145	Human	sec
106	11	2.0	239	4	ABR70310	AbR70310	Drosophill	179	11	2.0	440	6	ABR66763	AbR66763	Human	sec
107	11	2.0	249	4	ABR70394	AbR70394	Drosophill	180	11	2.0	440	6	ABO15976	AbO15976	Human	sec
108	11	2.0	249	5	ABJ04059	Abj04059	C parvum	181	11	2.0	440	6	ABO13682	AbO13682	Human	sec
109	11	2.0	252	8	ADP31485	Adp31485	Human sec	182	11	2.0	440	6	ABU57246	AbU57246	Human	PRO
110	11	2.0	255	4	ABB68072	Abb68072	Drosophill	183	11	2.0	440	6	ABU65585	AbU65585	Human	sec
111	11	2.0	267	8	ADX91941	Adx91941	Plant ful	184	11	2.0	440	6	ABO07433	AbO07433	Human	PRO
112	11	2.0	279	6	ABG71214	Abg71214	Rat p62 p	185	11	2.0	440	6	ABO03620	AbO03620	Human	sec
113	11	2.0	288	2	AA29082	Aay29082	T. gondii	186	11	2.0	440	6	ABR67068	AbR67068	Human	sec
114	11	2.0	288	2	AA29081	Aay29081	T. gondii	187	11	2.0	440	6	ABO15671	AbO15671	Human	sec
115	11	2.0	288	4	AA25553	Aau25553	T. gondii	188	11	2.0	440	6	ABU55952	AbU55952	Human	sec
116	11	2.0	288	4	AA25552	Aau25552	T. gondii	189	11	2.0	440	6	ABU65280	AbU65280	Human	PRO
117	11	2.0	288	7	ADG17391	Adg17391	T. gondii	190	11	2.0	440	6	ABU95225	AbU95225	Novel	hum
118	11	2.0	288	7	ADG17394	Adg17394	T. gondii	191	11	2.0	440	6	ABU71128	AbU71128	Human	PRO
119	11	2.0	308	8	AD254341	Ad254341	Bacterial	192	11	2.0	440	6	ABO07738	AbO07738	Human	PRO
120	11	2.0	311	8	ADN21547	Adn21547	Bacterial	193	11	2.0	440	6	ABR69979	AbR69979	Human	sec
121	11	2.0	313	7	ADL15085	Adl15085	Human CI	194	11	2.0	440	6	ABR69312	AbR69312	Human	sec
122	11	2.0	364	3	AA25586	Aab25586	Protein e	195	11	2.0	440	6	ABO01453	AbO01453	Human	PRO
123	11	2.0	364	3	ADA27058	Ada27058	Human nov	196	11	2.0	440	6	ABU81255	AbU81255	Human	PRO
124	11	2.0	364	8	AD25588	Ada25588	Novel hum	197	11	2.0	440	6	ABR60052	AbR60052	Human	sec
125	11	2.0	397	8	ABO84563	AbO84563	Mouse can	198	11	2.0	440	6	ABR67787	AbR67787	Human	sec
126	11	2.0	402	4	AA23691	Aam23691	Human EST	199	11	2.0	440	6	ABR65175	AbR65175	Human	sec
127	11	2.0	422	5	ABP74036	Abp74036	Candida a	200	11	2.0	440	6	ABR68397	AbR68397	Human	sec
128	11	2.0	423	3	AA45093	Aay45093	Mouse lym	201	11	2.0	440	6	ABR71809	AbR71809	Human	sec
129	11	2.0	423	9	ADM91678	Adm91678	Mouse LDC	202	11	2.0	440	6	ABU85289	AbU85289	Human	PRO
130	11	2.0	435	4	AB25785	Ab25785	Drosophill	203	11	2.0	440	6	ABU88879	AbU88879	Human	sec
131	11	2.0	435	5	ABP73423	Abp73423	Candida a	204	11	2.0	440	6	ABU83059	AbU83059	Human	sec
132	11	2.0	440	2	AA17830	Aay17830	Human PRO	205	11	2.0	440	6	ABU94915	AbU94915	Novel	hum
133	11	2.0	440	3	ABO1321	Aab01321	Human PRO	206	11	2.0	440	6	ABU90463	AbU90463	Human	sec
134	11	2.0	440	4	AAU29040	Aau29040	Human PRO	207	11	2.0	440	6	ABU83974	AbU83974	Human	sec
135	11	2.0	440	6	ABU58416	Abu58416	Human PRO	208	11	2.0	440	6	ABU93625	AbU93625	Novel	hum
136	11	2.0	440	6	ABU87964	Abu87964	Novel hum	209	11	2.0	440	6	ABR64870	AbR64870	Human	sec
137	11	2.0	440	6	ABU84279	Abu84279	Human sec	210	11	2.0	440	6	ABR68702	AbR68702	Human	sec
138	11	2.0	440	6	ABR66153	AbR66153	Human sec	211	11	2.0	440	6	ABO06518	AbO06518	Human	sec
139	11	2.0	440	6	ABR65543	AbR65543	Human sec	212	11	2.0	440	6	ABR99063	AbR99063	Human	sec
140	11	2.0	440	6	ABU99483	Abu99483	Human sec	213	11	2.0	440	6	ABU56311	AbU56311	Human	sec
141	11	2.0	440	6	ABU55930	Abu55930	Human sec	214	11	2.0	440	6	ABU56947	AbU56947	Human	PRO
142	11	2.0	440	6	ABU82722	Abu82722	Human PRO	215	11	2.0	440	6	ABU85899	AbU85899	Human	PRO
143	11	2.0	440	6	ABR89843	AbR89843	Novel hum	216	11	2.0	440	6	ABU82186	AbU82186	Human	sec
144	11	2.0	440	6	ABR68092	AbR68092	Human sec	217	11	2.0	440	6	ABU87197	AbU87197	Human	PRO
145	11	2.0	440	6	ABU96145	Abu96145	Novel hum	218	11	2.0	440	6	ABU83669	AbU83669	Human	sec
146	11	2.0	440	6	ABU92576	Abu92576	Human sec	219	11	2.0	440	6	ABO08043	AbO08043	Human	PRO
147	11	2.0	440	6	ABO08653	AbO08653	Human sec	220	11	2.0	440	6	ABU60351	AbU60351	Novel	hum
148	11	2.0	440	6	ABO02705	AbO02705	Human sec	221	11	2.0	440	6	ABU81754	AbU81754	Novel	hum
149	11	2.0	440	6	ABR74859	AbR74859	Human sec	222	11	2.0	440	6	ABR65918	AbR65918	Novel	hum
150	11	2.0	440	6	ABR94621	AbR94621	Human sec	223	11	2.0	440	6	ABR59747	AbR59747	Human	sec
151	11	2.0	440	6	ABU60240	Abu60240	Human PRO	224	11	2.0	440	6	ABU93935	AbU93935	Novel	hum
152	11	2.0	440	6	ABU85594	Abu85594	Human PRO	225	11	2.0	440	6	ABU99788	AbU99788	Novel	hum
153	11	2.0	440	6	ABU98754	Abu98754	Novel hum	226	11	2.0	440	6	ABR66458	AbR66458	Human	sec
154	11	2.0	440	6	ABU97969	Abu97969	Novel hum	227	11	2.0	440	6	ABR90876	AbR90876	Human	sec
155	11	2.0	440	6	ABU91675	Abu91675	Novel hum	228	11	2.0	440	6	ABU94303	AbU94303	Human	PRO
156	11	2.0	440	6	ABU89368	Abu89368	Human PRO	229	11	2.0	440	6	ABU79185	AbU79185	Human	PRO
157	11	2.0	440	6	ABU86209	Abu86209	Human sec	230	11	2.0	440	6	ABU86514	AbU86514	Human	sec
158	11	2.0	440	6	ABU67422	Abu67422	Human sec	231	11	2.0	440	6	ABU86819	AbU86819	Novel	hum
159	11	2.0	440	6	ABU80450	Abu80450	Human PRO	232	11	2.0	440	6	ABU94608	AbU94608	Human	PRO
160	11	2.0	440	6	ABR9368	AbR9368	Human sec	233	11	2.0	440	6	ABO04535	AbO04535	Human	PRO
161	11	2.0	440	6	ABR98758	AbR98758	Human sec	234	11	2.0	440	6	ABR70284	AbR70284	Human	sec
162	11	2.0	440	6	ABO16281	AbO16281	Human sec	235	11	2.0	440	6	ABU98449	AbU98449	Human	PRO
163	11	2.0	440	6	ABR92181	AbR92181	Human sec	236	11	2.0	440	6	ABR65848	AbR65848	Human	sec
164	11	2.0	440	6	ABO18822	AbO18822	Human sec	237	11	2.0	440	6	ABR64565	AbR64565	Human	sec
165	11	2.0	440	6	ABR78243	AbR78243	Human sec	238	11	2.0	440	6	ABU79480	AbU79480	Human	PRO
166	11	2.0	440	6	ABU64926	Abu64926	Human sec	239	11	2.0	440	6	ABU92881	AbU92881	Human	sec
167	11	2.0	440	6	ABU84979	Abu84979	Novel hum	240	11	2.0	440	6	ABU95840	AbU95840	Human	PRO
168	11	2.0	440	6	ABO00118	AbO00118	Novel hum	241	11	2.0	440	6	ABU91060	AbU91060	Novel	hum
169	11	2.0	440	6	ABO11450	AbO11450	Human sec	242	11	2.0	440	6	ABU90153	AbU90153	Novel	hum

243	11	2.0	440	6	ABO09568	AbO09568 Human sec	316	11	2.0	440	6	ABR75823	AbT75823 Human sec
244	11	2.0	440	6	ABO10840	AbO10840 Human sec	317	11	2.0	440	6	ABR71199	AbT71199 Human sec
245	11	2.0	440	6	ABR70894	AbT70894 Human sec	318	11	2.0	440	6	ABR93096	AbT93096 Human sec
246	11	2.0	440	6	ABU87502	AbU87502 Human PRO	319	11	2.0	440	6	ABR93401	AbT93401 Human sec
247	11	2.0	440	6	ABU91370	AbU91370 Human PRO	320	11	2.0	440	6	ABR87826	AbT87826 Human sec
248	11	2.0	440	6	ABU84584	AbU84584 Human sec	321	11	2.0	440	6	ABO27826	AbO27826 Human sec
249	11	2.0	440	6	ABR69674	AbR69674 Human sec	322	11	2.0	440	6	ABO32961	AbO32961 Human sec
250	11	2.0	440	6	ABU80051	AbU80051 Human PRO	323	11	2.0	440	6	ABO33170	AbO33170 Human PRO
251	11	2.0	440	6	ABU93320	AbU93320 Human PRO	324	11	2.0	440	6	ABM04858	AbM04858 Human PRO
252	11	2.0	440	6	ABO09873	AbO09873 Human sec	325	11	2.0	440	6	ABM08818	AbM08818 Human sec
253	11	2.0	440	6	ABO08958	AbO08958 Human sec	326	11	2.0	440	6	ABO36418	AbO36418 Human sec
254	11	2.0	440	6	ABU10526	AbU10526 Human sec	327	11	2.0	440	6	ABO35503	AbO35503 Human PRO
255	11	2.0	440	6	ABU11312	AbU11312 Human PRO	328	11	2.0	440	6	ABO39468	AbO39468 Human sec
256	11	2.0	440	6	ABU67131	AbU67131 Human PRO	329	11	2.0	440	6	ABM10343	AbM10343 Human sec
257	11	2.0	440	6	ABU95535	AbU95535 Human PRO	330	11	2.0	440	6	ABM11868	AbM11868 Human sec
258	11	2.0	440	6	ABU96744	AbU96744 Novel hum	331	11	2.0	440	6	ABO52014	AbO52014 Human PRO
259	11	2.0	440	6	ABR70589	AbT70589 Human sec	332	11	2.0	440	6	ABO52319	AbO52319 Human PRO
260	11	2.0	440	6	ABO04940	AbO04940 Novel hum	333	11	2.0	440	6	ABO32637	AbO32637 Human sec
261	11	2.0	440	6	ABO08348	AbO08348 Human sec	334	11	2.0	440	6	ABR97123	AbT97123 Human sec
262	11	2.0	440	6	ABO05555	AbO05555 Human sec	335	11	2.0	440	6	ABR86911	AbT86911 Human sec
263	11	2.0	440	6	ABR73944	AbT73944 Human sec	336	11	2.0	440	6	ABM10953	AbM10953 Human sec
264	11	2.0	440	6	ABR95336	AbT95336 Human sec	337	11	2.0	440	6	ABM28097	AbM28097 Human sec
265	11	2.0	440	6	ABR80833	AbT80833 Human sec	338	11	2.0	440	6	ABO32096	AbO32096 Human sec
266	11	2.0	440	6	ABR81138	AbT81138 Human sec	339	11	2.0	440	6	ABM15223	AbM15223 Human sec
267	11	2.0	440	6	ABM00834	AbM00834 Human sec	340	11	2.0	440	6	ABM06378	AbM06378 Human sec
268	11	2.0	440	6	ABR88436	AbT88436 Human sec	341	11	2.0	440	6	ABM04189	AbM04189 Human sec
269	11	2.0	440	6	ABM77257	AbM77257 Human sec	342	11	2.0	440	6	ABM22302	AbM22302 Human sec
270	11	2.0	440	6	ABO28741	AbO28741 Human sec	343	11	2.0	440	6	ABM07598	AbM07598 Human sec
271	11	2.0	440	6	ABO31486	AbO31486 Human sec	344	11	2.0	440	6	ABO40688	AbO40688 Human sec
272	11	2.0	440	6	ABM07903	AbM07903 Human sec	345	11	2.0	440	6	ABM35335	AbM35335 Human sec
273	11	2.0	440	6	ABO40383	AbO40383 Human sec	346	11	2.0	440	6	ABM33098	AbM33098 Human sec
274	11	2.0	440	6	ABO35808	AbO35808 Human PRO	347	11	2.0	440	6	ABO52624	AbO52624 Human PRO
275	11	2.0	440	6	ABO43947	AbO43947 Human PRO	348	11	2.0	440	6	ABO50184	AbO50184 Human sec
276	11	2.0	440	6	ADA77786	Ada77786 Human sec	349	11	2.0	440	6	ABU99178	AbU99178 Human sec
277	11	2.0	440	6	ABM24742	Abm24742 Human sec	350	11	2.0	440	6	ABO44230	AbO44230 Human sec
278	11	2.0	440	6	ABR90266	AbR90266 Human sec	351	11	2.0	440	6	ABO05860	AbO05860 Human sec
279	11	2.0	440	6	ABM17180	AbM17180 Human sec	352	11	2.0	440	6	ABM18400	AbM18400 Human sec
280	11	2.0	440	6	ABR94926	AbT94926 Human sec	353	11	2.0	440	6	ABR97428	AbT97428 Human sec
281	11	2.0	440	6	ABR94926	AbT94926 Human sec	354	11	2.0	440	6	ABR80528	AbT80528 Human sec
282	11	2.0	440	6	ABR95231	AbT95231 Human sec	355	11	2.0	440	6	ABM01139	AbM01139 Human sec
283	11	2.0	440	6	ABO21469	AbO21469 Human sec	356	11	2.0	440	6	ABR88741	AbT88741 Human sec
284	11	2.0	440	6	ABR97733	AbT97733 Human sec	357	11	2.0	440	6	ABM13393	AbM13393 Human sec
285	11	2.0	440	6	ABR87521	AbT87521 Human sec	358	11	2.0	440	6	ABM20777	AbM20777 Human sec
286	11	2.0	440	6	ABM77562	AbM77562 Human sec	359	11	2.0	440	6	ABO41908	AbO41908 Human sec
287	11	2.0	440	6	ABM27792	AbM27792 Human sec	360	11	2.0	440	6	ABO42518	AbO42518 Human sec
288	11	2.0	440	6	ABM06073	Abm06073 Human sec	361	11	2.0	440	6	ABM10038	AbM10038 Human sec
289	11	2.0	440	6	ABM03579	Abm03579 Human sec	362	11	2.0	440	6	ABO38553	AbO38553 Human sec
290	11	2.0	440	6	ABM35030	Abm35030 Human sec	363	11	2.0	440	6	ABM32793	AbM32793 Human sec
291	11	2.0	440	6	ABM26267	Abm26267 Human sec	364	11	2.0	440	6	ABM22607	AbM22607 Human sec
292	11	2.0	440	6	ABO48049	AbO48049 Human sec	365	11	2.0	440	6	ABM74818	AbM74818 Human sec
293	11	2.0	440	6	ABR92791	AbT92791 Human sec	366	11	2.0	440	6	ADA79578	Ada79578 Human sec
294	11	2.0	440	6	ABO24552	AbO24552 Human sec	367	11	2.0	440	6	ABR96208	AbT96208 Human sec
295	11	2.0	440	6	ABM11563	Abm11563 Human sec	368	11	2.0	440	6	ABM02359	AbM02359 Human sec
296	11	2.0	440	6	ABM02664	Abm02664 Human sec	369	11	2.0	440	6	ABR86301	AbT86301 Human sec
297	11	2.0	440	6	ABM15960	Abm15960 Human sec	370	11	2.0	440	6	ABR86606	AbT86606 Human sec
298	11	2.0	440	6	ABO27521	AbO27521 Human sec	371	11	2.0	440	6	ABM16570	AbM16570 Human sec
299	11	2.0	440	6	ABM29012	Abm29012 Human sec	372	11	2.0	440	6	ABM29622	AbM29622 Human sec
300	11	2.0	440	6	ABM06988	Abm06988 Human sec	373	11	2.0	440	6	ABO29046	AbO29046 Human sec
301	11	2.0	440	6	ABM21082	Abm21082 Human sec	374	11	2.0	440	6	ABM23827	AbM23827 Human sec
302	11	2.0	440	6	ABM09428	Abm09428 Human sec	375	11	2.0	440	6	ABM23217	AbM23217 Human sec
303	11	2.0	440	6	ABO41298	AbO41298 Human sec	376	11	2.0	440	6	ABM21997	AbM21997 Human sec
304	11	2.0	440	6	ABO36113	AbO36113 Human PRO	377	11	2.0	440	6	ABO37638	AbO37638 Human sec
305	11	2.0	440	6	ABO43642	AbO43642 Human PRO	378	11	2.0	440	6	ABM28402	AbM28402 Human sec
306	11	2.0	440	6	ABM76342	Abm76342 Human sec	379	11	2.0	440	6	ABM28707	AbM28707 Human sec
307	11	2.0	440	6	ABM76038	Abm76038 Human sec	380	11	2.0	440	6	ABM66351	AbM66351 Human sec
308	11	2.0	440	6	ABM25657	Abm25657 Human sec	381	11	2.0	440	6	ABM75733	AbM75733 Human sec
309	11	2.0	440	6	ABM25962	Abm25962 Human sec	382	11	2.0	440	6	ABM34013	AbM34013 Human sec
310	11	2.0	440	6	ABO03315	AbO03315 Human sec	383	11	2.0	440	6	ABM34318	AbM34318 Human sec
311	11	2.0	440	6	ABO02400	AbO02400 Human sec	384	11	2.0	440	6	ABO20249	AbO20249 Human sec
312	11	2.0	440	6	ABR90571	AbT90571 Human sec	385	11	2.0	440	6	ABO21164	AbO21164 Human sec
313	11	2.0	440	6	ABR73639	AbT73639 Human sec	386	11	2.0	440	6	ABO22079	AbO22079 Human sec
314	11	2.0	440	6	ABO16891	AbO16891 Human sec	387	11	2.0	440	6	ABR96513	AbT96513 Human sec
315	11	2.0	440	6	ABR94316	AbT94316 Human sec	388	11	2.0	440	6	ABR85691	AbT85691 Human sec

389	11	2.0	440	6	ABR99673	Abt99673	Human	sec	462	11	2.0	440	6	ABM05463	Abm05463	Human	sec
390	11	2.0	440	6	ABM00529	Abm00529	Human	sec	463	11	2.0	440	6	ABM15528	Abm15528	Human	sec
391	11	2.0	440	6	ABM00224	Abm00224	Human	sec	464	11	2.0	440	6	ABM08513	Abm08513	Human	sec
392	11	2.0	440	6	ABO29656	AbO29656	Human	sec	465	11	2.0	440	6	ABO42213	AbO42213	Human	sec
393	11	2.0	440	6	ABM23522	Abm23522	Human	sec	466	11	2.0	440	6	ABO37943	AbO37943	Human	sec
394	11	2.0	440	6	ABM29317	Abm29317	Human	sec	467	11	2.0	440	6	ABO45853	AbO45853	Human	sec
395	11	2.0	440	6	ABO38248	AbO38248	Human	sec	468	11	2.0	440	6	ABM66556	ABM66556	Human	sec
396	11	2.0	440	6	ABO45548	AbO45548	Human	sec	469	11	2.0	440	6	ABM66556	ABM66556	Human	sec
397	11	2.0	440	6	ABO20472	Abm20472	Human	PRO	470	11	2.0	440	6	ABM19557	Abm19557	Human	sec
398	11	2.0	440	6	ADA81305	Ada81305	Human	sec	471	11	2.0	440	6	ABO49269	ABO49269	Human	sec
399	11	2.0	440	6	ABO16586	AbO16586	Human	sec	472	11	2.0	440	6	ABO49574	ABO49574	Human	sec
400	11	2.0	440	6	ABO18212	AbO18212	Human	sec	473	11	2.0	440	6	ADA78398	Ada78398	Human	sec
401	11	2.0	440	6	ABO22639	AbO22639	Human	PRO	474	11	2.0	440	6	ABR88131	ABR88131	Human	sec
402	11	2.0	440	6	ABO22944	AbO22944	Human	PRO	475	11	2.0	440	6	ABM26877	ABM26877	Human	sec
403	11	2.0	440	6	ABR92486	AbR92486	Human	sec	476	11	2.0	440	6	ABM03274	ABM03274	Human	sec
404	11	2.0	440	6	ABR81443	AbR81443	Human	sec	477	11	2.0	440	7	ABO39773	ABO39773	Human	sec
405	11	2.0	440	6	ABM77867	Abm77867	Human	sec	478	11	2.0	440	7	ABO49879	ABO49879	Human	sec
406	11	2.0	440	6	ABR89656	AbR89656	Human	sec	479	11	2.0	440	7	ABO50794	ABO50794	Human	sec
407	11	2.0	440	6	ABM26572	Abm26572	Human	sec	480	11	2.0	440	7	ABO05250	ABO05250	Human	sec
408	11	2.0	440	6	ABM13698	Abm13698	Human	sec	481	11	2.0	440	7	ABR74554	ABR74554	Human	sec
409	11	2.0	440	6	ABO38436	AbO38436	Human	sec	482	11	2.0	440	7	ABR77033	ABR77033	Human	sec
410	11	2.0	440	6	ABO30256	AbO30256	Human	sec	483	11	2.0	440	7	ABM17790	ABM17790	Human	sec
411	11	2.0	440	6	ABM07293	Abm07293	Human	sec	484	11	2.0	440	7	ABR95841	ABR95841	Human	sec
412	11	2.0	440	6	ABM03884	Abm03884	Human	sec	485	11	2.0	440	7	ABO21774	ABO21774	Human	sec
413	11	2.0	440	6	ABO37028	AbO37028	Human	sec	486	11	2.0	440	7	ABO19944	ABO19944	Human	sec
414	11	2.0	440	6	ABO41603	AbO41603	Human	sec	487	11	2.0	440	7	ABO24247	ABO24247	Human	sec
415	11	2.0	440	6	ABO35198	AbO35198	Human	PRO	488	11	2.0	440	7	ABR85996	ABR85996	Human	sec
416	11	2.0	440	6	ABM25047	Abm25047	Human	sec	489	11	2.0	440	7	ABM10648	ABM10648	Human	sec
417	11	2.0	440	6	ABO47439	AbO47439	Human	sec	490	11	2.0	440	7	ABM76647	ABM76647	Human	sec
418	11	2.0	440	6	ABO47744	AbO47744	Human	sec	491	11	2.0	440	7	ABR89351	ABR89351	Human	sec
419	11	2.0	440	6	ABO48354	AbO48354	Human	sec	492	11	2.0	440	7	ABM12478	ABM12478	Human	sec
420	11	2.0	440	6	ABO51404	AbO51404	Human	PRO	493	11	2.0	440	7	ABM05768	ABM05768	Human	sec
421	11	2.0	440	6	ABO51709	AbO51709	Human	PRO	494	11	2.0	440	7	ABO34893	ABO34893	Human	PRO
422	11	2.0	440	6	ABO50489	AbO50489	Human	sec	495	11	2.0	440	7	ABM02969	ABM02969	Human	sec
423	11	2.0	440	6	ABR79613	AbR79613	Human	sec	496	11	2.0	440	7	ABM18947	ABM18947	Human	sec
424	11	2.0	440	6	ABM16875	Abm16875	Human	sec	497	11	2.0	440	7	ABM19252	ABM19252	Human	sec
425	11	2.0	440	6	ABO17907	AbO17907	Human	sec	498	11	2.0	440	7	ABO46463	ABO46463	Human	PRO
426	11	2.0	440	6	ABO20859	AbO20859	Human	sec	499	11	2.0	440	7	ABO48964	ABO48964	Human	sec
427	11	2.0	440	6	ABR96818	AbR96818	Human	sec	500	11	2.0	440	7	ABR69007	ABR69007	Human	sec
428	11	2.0	440	6	ABM12173	Abm12173	Human	sec	501	11	2.0	440	7	ABR89046	ABR89046	Human	sec
429	11	2.0	440	6	ABM16265	Abm16265	Human	sec	502	11	2.0	440	7	ABR72419	ABR72419	Human	sec
430	11	2.0	440	6	ABM24132	Abm24132	Human	sec	503	11	2.0	440	7	ABR74249	ABR74249	Human	sec
431	11	2.0	440	6	ABM14613	Abm14613	Human	sec	504	11	2.0	440	7	ABO18517	ABO18517	Human	sec
432	11	2.0	440	6	ABM04494	Abm04494	Human	sec	505	11	2.0	440	7	ABR80223	ABR80223	Human	sec
433	11	2.0	440	6	ABM06683	Abm06683	Human	sec	506	11	2.0	440	7	ABM01444	ABM01444	Human	sec
434	11	2.0	440	6	ABM09123	Abm09123	Human	sec	507	11	2.0	440	7	ABM02054	ABM02054	Human	sec
435	11	2.0	440	6	ABO39163	AbO39163	Human	sec	508	11	2.0	440	7	ABR87216	ABR87216	Human	sec
436	11	2.0	440	6	ABM75428	Abm75428	Human	sec	509	11	2.0	440	7	ABM12783	ABM12783	Human	sec
437	11	2.0	440	6	ABM25352	Abm25352	Human	sec	510	11	2.0	440	7	ABM30537	ABM30537	Human	sec
438	11	2.0	440	6	ABM19862	Abm19862	Human	sec	511	11	2.0	440	7	ABM24437	ABM24437	Human	sec
439	11	2.0	440	6	ABO46768	AbO46768	Human	PRO	512	11	2.0	440	7	ABO29351	ABO29351	Human	sec
440	11	2.0	440	6	ABO47073	AbO47073	Human	PRO	513	11	2.0	440	7	ABO31181	ABO31181	Human	sec
441	11	2.0	440	6	ADA83103	Ada83103	Human	sec	514	11	2.0	440	7	ABM14308	ABM14308	Human	sec
442	11	2.0	440	6	ABR71504	AbR71504	Human	sec	515	11	2.0	440	7	ABM09733	ABM09733	Human	sec
443	11	2.0	440	6	ABR72114	AbR72114	Human	sec	516	11	2.0	440	7	ABO38858	ABO38858	Human	sec
444	11	2.0	440	6	ABR98453	AbR98453	Human	sec	517	11	2.0	440	7	ABM34623	ABM34623	Human	sec
445	11	2.0	440	6	ABO06823	AbO06823	Human	sec	518	11	2.0	440	7	ABO51099	ABO51099	Human	sec
446	11	2.0	440	6	ABR84776	AbR84776	Human	sec	519	11	2.0	440	7	ABO03925	ABO03925	Human	sec
447	11	2.0	440	6	ABR73334	AbR73334	Human	sec	520	11	2.0	440	7	ABO10395	ABO10395	Human	PRO
448	11	2.0	440	6	ABR76428	AbR76428	Human	sec	521	11	2.0	440	7	ABR77638	ABR77638	Human	sec
449	11	2.0	440	6	ABR73029	AbR73029	Human	sec	522	11	2.0	440	7	ABR78848	ABR78848	Human	sec
450	11	2.0	440	6	ABM18095	Abm18095	Human	sec	523	11	2.0	440	7	ABO23942	ABO23942	Human	sec
451	11	2.0	440	6	ABO20554	AbO20554	Human	sec	524	11	2.0	440	7	ABR93706	ABR93706	Human	sec
452	11	2.0	440	6	ABO25297	AbO25297	Human	PRO	525	11	2.0	440	7	ABM01749	ABM01749	Human	sec
453	11	2.0	440	6	ABO25602	AbO25602	Human	PRO	526	11	2.0	440	7	ABM78172	ABM78172	Human	sec
454	11	2.0	440	6	ABR94011	AbR94011	Human	sec	527	11	2.0	440	7	ABR89961	ABR89961	Human	sec
455	11	2.0	440	6	ABR79918	AbR79918	Human	sec	528	11	2.0	440	7	ABM27487	ABM27487	Human	sec
456	11	2.0	440	6	ABM11258	Abm11258	Human	sec	529	11	2.0	440	7	ABM13088	ABM13088	Human	sec
457	11	2.0	440	6	ABO32865	AbO32865	Human	PRO	530	11	2.0	440	7	ABO31791	ABO31791	Human	sec
458	11	2.0	440	6	ABO30571	AbO30571	Human	sec	531	11	2.0	440	7	ABM14003	ABM14003	Human	sec
459	11	2.0	440	6	ABO30876	AbO30876	Human	sec	532	11	2.0	440	7	ABM08208	ABM08208	Human	sec
460	11	2.0	440	6	ABM27182	Abm27182	Human	sec	533	11	2.0	440	7	ABO40078	ABO40078	Human	sec
461	11	2.0	440	6	ABM29927	Abm29927	Human	sec	534	11	2.0	440	7	ABM74513	ABM74513	Human	sec

535	1.1	2.0	440	7	ABM33708	Abm33708 Human sec	608	1.1	2.0	442	6	ADA27144	Ada27144 Human nov
536	1.1	2.0	440	7	ABM20167	Abm20167 Human sec	609	1.1	2.0	442	7	ADBE1605	Adbe1605 Human pro
537	1.1	2.0	440	7	ABO48659	AbO48659 Human sec	610	1.1	2.0	442	7	ADBE1608	Adbe1608 Human pro
538	1.1	2.0	440	7	ABR72724	AbR72724 Human sec	611	1.1	2.0	442	7	ADBE54238	Adbe54238 Human pro
539	1.1	2.0	440	7	ABO15366	AbO15366 Human sec	612	1.1	2.0	442	8	ADBE6685	Adbe6685 Novel hum
540	1.1	2.0	440	7	ABR85081	AbR85081 Human sec	613	1.1	2.0	442	8	ADU06700	AdU06700 Novel bro
541	1.1	2.0	440	7	ABO15061	AbO15061 Human sec	614	1.1	2.0	442	9	ADW91676	AdW91676 Human LDC
542	1.1	2.0	440	7	ABO17196	AbO17196 Human sec	615	1.1	2.0	442	9	ADZ70675	AdZ70675 Human pro
543	1.1	2.0	440	7	ABM17485	Abm17485 Human sec	616	1.1	2.0	442	9	AEAS2682	AEAS2682 Human TSL
544	1.1	2.0	440	7	ABR85386	AbR85386 Human sec	617	1.1	2.0	460	8	ADY06832	ADY06832 Plant ful
545	1.1	2.0	440	7	ABM76952	Abm76952 Human sec	618	1.1	2.0	474	4	ABBE63408	ABBE63408 Drosophi
546	1.1	2.0	440	7	ABO28131	AbO28131 Human sec	619	1.1	2.0	496	8	ADP04787	ADP04787 Sea squit
547	1.1	2.0	440	7	ABM22912	Abm22912 Human sec	620	1.1	2.0	498	5	ABP73752	ABP73752 Candida a
548	1.1	2.0	440	7	ABM30232	Abm30232 Human sec	621	1.1	2.0	503	3	AA631844	AA631844 Arabidops
549	1.1	2.0	440	7	ABM21692	Abm21692 Human sec	622	1.1	2.0	512	3	AA631843	AA631843 Arabidops
550	1.1	2.0	440	7	ABM21387	Abm21387 Human sec	623	1.1	2.0	517	2	AAW04558	AAW04558 Carnation
551	1.1	2.0	440	7	ABM14918	Abm14918 Human sec	624	1.1	2.0	519	7	ADD55824	ADD55824 Thalecres
552	1.1	2.0	440	7	ABO40993	AbO40993 Human sec	625	1.1	2.0	519	8	ADOO2205	ADOO2205 Thalecres
553	1.1	2.0	440	7	ABO36723	AbO36723 Human sec	626	1.1	2.0	519	8	ADT06983	ADT06983 Arabidops
554	1.1	2.0	440	7	ABO37333	AbO37333 Human sec	627	1.1	2.0	520	4	ABBE63354	ABBE63354 Drosophi
555	1.1	2.0	440	7	ABM75123	Abm75123 Human sec	628	1.1	2.0	522	4	AA684964	AA684964 Shrimp wh
556	1.1	2.0	440	7	ABM33403	Abm33403 Human sec	629	1.1	2.0	523	7	ADC87177	ADC87177 Human GPC
557	1.1	2.0	440	7	ABO46158	AbO46158 Human sec	630	1.1	2.0	523	8	ADX95781	ADX95781 Plant ful
558	1.1	2.0	440	7	ADA82469	AdA82469 Human sec	631	1.1	2.0	525	6	ABG71213	ABG71213 Full- leng
559	1.1	2.0	440	7	ABM31757	Abm31757 Human sec	632	1.1	2.0	525	7	ADBE57923	ADBE57923 Rat Prote
560	1.1	2.0	440	7	ABM31147	Abm31147 Human sec	633	1.1	2.0	525	7	ADBE57925	ADBE57925 Rat Prote
561	1.1	2.0	440	7	ADBE85777	AdBE85777 Human sec	634	1.1	2.0	559	5	AAW48982	AAW48982 Orphan in
562	1.1	2.0	440	7	ABM32062	Abm32062 Human sec	635	1.1	2.0	566	8	ADRI4944	ADRI4944 C. albica
563	1.1	2.0	440	7	ABM32367	Abm32367 Human sec	636	1.1	2.0	571	9	AE887650	AE887650 Candida a
564	1.1	2.0	440	7	ABM14152	Abm14152 Human sec	637	1.1	2.0	681	5	AE856727	AE856727 C. elegan
565	1.1	2.0	440	7	ABM30842	Abm30842 Human sec	638	1.1	2.0	694	5	ABBE63094	ABBE63094 FLO11 gen
566	1.1	2.0	440	7	ADC25817	AdC25817 Human sec	639	1.1	2.0	707	7	ADD25157	ADD25157 Fertility
567	1.1	2.0	440	7	ADC25575	AdC25575 Human sec	640	1.1	2.0	707	8	ADN61172	ADN61172 Radish nu
568	1.1	2.0	440	7	ADC25696	AdC25696 Human sec	641	1.1	2.0	725	4	ABBE68684	ABBE68684 Drosophi
569	1.1	2.0	440	7	ADD05507	AdD05507 Human sec	642	1.1	2.0	730	4	ABG27231	ABG27231 Novel hum
570	1.1	2.0	440	7	ADG02502	AdG02502 Novel hum	643	1.1	2.0	746	5	ABBS59201	ABBS59201 Drosophi
571	1.1	2.0	440	7	ADG01209	AdG01209 Novel hum	644	1.1	2.0	769	5	ABP73280	ABP73280 Candida a
572	1.1	2.0	440	7	ADP95384	AdP95384 Novel hum	645	1.1	2.0	793	9	AE887653	AE887653 Dictyoste
573	1.1	2.0	440	7	ADG12199	AdG12199 Novel hum	646	1.1	2.0	847	2	AAI1988	AAI1988 Dictyoste
574	1.1	2.0	440	7	ADH27481	AdH27481 Human sec	647	1.1	2.0	945	2	AAV06119	AAV06119 Human CIT
575	1.1	2.0	440	7	ADH08859	AdH08859 Human PRO	648	1.1	2.0	1002	3	AAV79166	AAV79166 Pneumocys
576	1.1	2.0	440	7	ADL32640	AdL32640 Novel hum	649	1.1	2.0	1057	4	ABBE65440	ABBE65440 Drosophi
577	1.1	2.0	440	7	ADM30174	AdM30174 Novel hum	650	1.1	2.0	1077	5	ABP74094	ABP74094 Candida a
578	1.1	2.0	440	8	ADE74171	AdE74171 Human sec	651	1.1	2.0	1088	4	ABBE67939	ABBE67939 Drosophi
579	1.1	2.0	440	8	ADE71530	AdE71530 Human sec	652	1.1	2.0	1088	8	ADOO8108	ADOO8108 FLY POLYP
580	1.1	2.0	440	8	ADE74783	AdE74783 Human sec	653	1.1	2.0	1128	6	ADA15725	ADA15725 C. elegan
581	1.1	2.0	440	8	ADP95996	AdP95996 Novel hum	654	1.1	2.0	1128	10	AER43022	AER43022 Nematode
582	1.1	2.0	440	8	ADG04267	AdG04267 Novel hum	655	1.1	2.0	1137	4	ABBS8536	ABBS8536 Drosophi
583	1.1	2.0	440	8	ADG00427	AdG00427 Novel hum	656	1.1	2.0	1271	6	ABU35651	ABU35651 Protein e
584	1.1	2.0	440	8	ADG63473	AdG63473 Human sec	657	1.1	2.0	1313	7	ADJ70374	ADJ70374 Human hea
585	1.1	2.0	440	8	ADG82683	AdG82683 Human PRO	658	1.1	2.0	1428	4	ABBE60377	ABBE60377 Drosophi
586	1.1	2.0	440	8	ADH25964	AdH25964 Novel hum	659	1.1	2.0	1468	4	ABBE65329	ABBE65329 Drosophi
587	1.1	2.0	440	8	ADH32933	AdH32933 Human PRO	660	1.1	2.0	1616	8	ADP22958	ADP22958 PRO POLYP
588	1.1	2.0	440	8	ADH43202	AdH43202 Human sec	661	1.1	2.0	1682	8	ADRI4131	ADRI4131 Human NF-
589	1.1	2.0	440	8	ADU54672	AdU54672 Human PRO	662	1.1	2.0	1721	2	AAW48299	AAW48299 Cryptospo
590	1.1	2.0	440	8	ADU64443	AdU64443 Human PRO	663	1.1	2.0	1721	5	AAH11727	AAH11727 Portlion o
591	1.1	2.0	440	8	ADM31339	AdM31339 Novel hum	664	1.1	2.0	1721	5	ABU04045	ABU04045 C parvum
592	1.1	2.0	440	8	ADM36386	AdM36386 Novel hum	665	1.1	2.0	1837	5	AAH11726	AAH11726 Cryptospo
593	1.1	2.0	440	8	ADM40191	AdM40191 Novel hum	666	1.1	2.0	1837	5	ABJ04064	ABJ04064 C parvum
594	1.1	2.0	440	8	ADN00440	AdN00440 Human sec	667	1.1	2.0	2123	8	ADP30657	ADP30657 Human sec
595	1.1	2.0	440	8	ADN37799	AdN37799 Novel hum	668	1.1	2.0	2123	8	ADP30564	ADP30564 Human sec
596	1.1	2.0	440	8	ADU25364	AdU25364 Human sec	669	1.1	2.0	2296	4	ABBS9050	ABBS9050 Drosophi
597	1.1	2.0	440	9	ADY39527	AdY39527 Human CRT	670	1.1	2.0	4544	8	ADP25443	ADP25443 Plasmodiu
598	1.1	2.0	440	9	ADY73808	AdY73808 Human PRO	671	1.1	1.8	10	4	AAW43253	AAW43253 Mycoplae
599	1.1	2.0	440	9	AED50029	AEd50029 Novel hum	672	1.1	1.8	10	4	AAW43243	AAW43243 Mycoplae
600	1.1	2.0	442	3	AAE25619	AAe25619 Protein e	673	1.1	1.8	10	4	AAW43227	AAW43227 Mycoplae
601	1.1	2.0	442	3	AAE94341	AAe94341 Human cel	674	1.1	1.8	10	4	AAW43251	AAW43251 Mycoplae
602	1.1	2.0	442	3	AAE45092	AAe45092 Human lym	675	1.1	1.8	10	4	AAW43217	AAW43217 Mycoplae
603	1.1	2.0	442	4	AAE62210	AAe62210 Drosophi	676	1.1	1.8	10	4	AAW43225	AAW43225 Mycoplae
604	1.1	2.0	442	5	AAE19887	AAe19887 Human tum	677	1.1	1.8	10	4	AAW43241	AAW43241 Mycoplae
605	1.1	2.0	442	5	ABP62825	ABp62825 Human pol	678	1.1	1.8	10	4	AEF53945	AEF53945 Size bran
606	1.1	2.0	442	6	ABO07196	AbO07196 Human p53	679	1.1	1.8	10	10	AEF53945	AEF53945 Peptide S
607	1.1	2.0	442	6	ABO07231	AbO07231 Human p53	680	1.1	1.8	11	10	AEF53955	AEF53955 Peptide S

681	10	1.8	20	6	ABU62603	Abu62603 Hypoetheti-	754	9	1.6	63	4	AAW21396	Aam21396 Peptide #
682	10	1.8	29	10	AE837119	Ae837119 Human ser	755	9	1.6	63	4	ABB43733	Abb43733 Peptide #
683	10	1.8	50	4	AAW84845	Aaw84845 Human imm	756	9	1.6	63	4	AAW37631	Aaw37631 Peptide #
684	10	1.8	107	3	AAZ20576	Aaz20576 Intimin C	757	9	1.6	63	4	ABB26679	Abb26679 Protein #
685	10	1.8	115	4	ABG26497	Abg26497 Novel hum	758	9	1.6	63	4	AAW77470	Aaw77470 Human Don
686	10	1.8	127	3	AAW07769	Aaw07769 Arabidops	759	9	1.6	63	4	AAW64695	Aaw64695 Human bra
687	10	1.8	133	10	AEF29277	Aef29277 Lead Cere	760	9	1.6	63	4	ABG59106	Abg59106 Human liv
688	10	1.8	158	3	AAW07768	Aaw07768 Arabidops	761	9	1.6	63	5	ABG46488	Abg46488 Human pep
689	10	1.8	158	8	ADW87818	Adw87818 plant ful	762	9	1.6	74	4	AAW73978	Aaw73978 Human col
690	10	1.8	161	10	AE826473	Ae826473 ORF0826 i	763	9	1.6	93	4	ABB68547	Abb68547 Drosophil
691	10	1.8	168	8	ADP31621	Adp31621 Human sec	764	9	1.6	102	3	AAW02285	Aaw02285 Human sec
692	10	1.8	186	3	AAW28749	Aaw28749 Arabidops	765	9	1.6	136	5	ABB99954	Abb99954 Human pro
693	10	1.8	187	3	AAW28748	Aaw28748 Arabidops	766	9	1.6	142	3	AAW37454	Aaw37454 Arabidops
694	10	1.8	195	4	ABB69342	Abb69342 Drosophil	767	9	1.6	161	7	AEW87495	Aew87495 Rice abio
695	10	1.8	202	3	AAW28747	Aaw28747 Arabidops	768	9	1.6	168	8	ADP31099	Adp31099 Human sec
696	10	1.8	210	7	ABW82140	Abw82140 Pseudomon	769	9	1.6	169	8	ADW74706	Adw74706 Plant ful
697	10	1.8	211	2	AAW81141	Aaw81141 A Histpla	770	9	1.6	176	3	AAW37453	Aaw37453 Arabidops
698	10	1.8	253	6	ADA15489	Ada15489 A. thalia	771	9	1.6	184	10	AEF28989	Aef28989 Lead Cere
699	10	1.8	253	4	ADW01605	Adw01605 Thalecres	772	9	1.6	186	4	ABB69324	Abb69324 Drosophil
700	10	1.8	254	8	ABB69981	Abb69981 Drosophil	773	9	1.6	202	5	AAW37452	Aaw37452 Laccococc
701	10	1.8	256	8	ADW67355	Adw67355 plant ful	774	9	1.6	224	4	ABB71117	Abb71117 Drosophil
702	10	1.8	259	10	AEF28747	Aef28747 Lead Cere	775	9	1.6	224	4	ABB71117	Abb71117 Drosophil
703	10	1.8	264	8	ADW02303	Adw02303 Novel hum	776	9	1.6	241	4	ABB70942	Abb70942 Drosophil
704	10	1.8	307	4	ABB65879	Abb65879 Drosophil	777	9	1.6	273	8	ADW62727	Adw62727 Transcript
705	10	1.8	335	7	ADW31825	Adw31825 plant (A.	778	9	1.6	296	8	ABW60442	Abw60442 Human gen
706	10	1.8	335	7	ADD30190	Ad30190 plant yle	779	9	1.6	306	4	ABB65424	Abb65424 Drosophil
707	10	1.8	335	8	ADW43603	Adw43603 plant tra	780	9	1.6	323	8	ADW69064	Adw69064 Plant ful
708	10	1.8	335	8	ADW02215	Adw02215 Thalecres	781	9	1.6	348	7	ABW85833	Abw85833 Human pro
709	10	1.8	335	8	ADW03407	Adw03407 Thalecres	782	9	1.6	378	9	AEC74778	Aec74778 Aspergill
710	10	1.8	335	8	AEW26677	Aew26677 Strees to	783	9	1.6	389	4	ABB68044	Abb68044 Drosophil
711	10	1.8	370	5	ADW41469	Adw41469 Human CD-	784	9	1.6	397	9	AEC74753	Aec74753 Aspergill
712	10	1.8	394	3	AAW54255	Aaw54255 Arabidops	785	9	1.6	490	10	AEF28861	Aef28861 Lead Cere
713	10	1.8	395	3	AAW54254	Aaw54254 Arabidops	786	9	1.6	492	8	ADW13964	Adw13964 Rat methy
714	10	1.8	409	3	AAW54253	Aaw54253 Arabidops	787	9	1.6	494	4	ABB70094	Abb70094 Drosophil
715	10	1.8	491	5	ABP73324	Abp73324 Candida a	788	9	1.6	499	8	ADW70635	Adw70635 Plant ful
716	10	1.8	528	6	ADW01785	Adw01785 BFLP0169	789	9	1.6	506	4	ABB58157	Abb58157 Drosophil
717	10	1.8	528	6	ABU45017	Abu45017 Protein e	790	9	1.6	526	3	AAW84906	Aaw84906 A human p
718	10	1.8	533	4	ABB86965	Abb86965 D. melano	791	9	1.6	526	4	AAW78533	Aaw78533 Human pro
719	10	1.8	579	4	ABB71231	Abb71231 Drosophil	792	9	1.6	526	4	AAW94371	Aaw94371 Human pro
720	10	1.8	625	8	ADW92996	Adw92996 plant ful	793	9	1.6	526	6	ABW69606	Abw69606 Human NF-
721	10	1.8	713	8	ADW97079	Adw97079 C. albica	794	9	1.6	526	7	ADW54820	Adw54820 Human pro
722	10	1.8	788	4	ABB60583	Abb60583 Drosophil	795	9	1.6	526	7	ADW54820	Adw54820 Human pro
723	10	1.8	788	4	AAW38171	Aaw38171 Fruit fly	796	9	1.6	526	7	ADW54832	Adw54832 Human pro
724	10	1.8	842	5	ABP73474	Abp73474 Candida a	797	9	1.6	526	7	ADW54828	Adw54828 Human pro
725	10	1.8	1009	3	AAW79167	Aaw79167 Pneumocys	798	9	1.6	526	8	ADW25161	Adw25161 PRO polyP
726	10	1.8	1014	5	AAW17503	Aaw17503 C elegans	799	9	1.6	526	8	ADW88055	Adw88055 Human pro
727	10	1.8	1017	3	AAW79170	Aaw79170 Pneumocys	800	9	1.6	526	8	ADW83165	Adw83165 Human pro
728	10	1.8	1023	3	AAW79171	Aaw79171 Pneumocys	801	9	1.6	527	7	ADW54832	Adw54832 Rat prote
729	10	1.8	1027	3	AAW79168	Aaw79168 Pneumocys	802	9	1.6	527	7	ADW54830	Adw54830 Rat prote
730	10	1.8	1029	3	AAW79167	Aaw79167 Pneumocys	803	9	1.6	527	7	ADW54826	Adw54826 Rat prote
731	10	1.8	1111	8	ADW01783	Adw01783 BFLP0169	804	9	1.6	527	7	ADW54818	Adw54818 Rat prote
732	10	1.8	1307	4	ABB65464	Abb65464 Drosophil	805	9	1.6	527	9	ADW83188	Adw83188 Rat cycli
733	10	1.8	1373	5	ABP73292	Abp73292 Candida a	806	9	1.6	533	8	ADW31279	Adw31279 Human sec
734	10	1.8	1419	5	ABP69842	Abb69842 Human pol	807	9	1.6	533	5	ABW70759	Abw70759 T. mariti
735	10	1.8	1448	7	ADW61912	Adw61912 Human hea	808	9	1.6	554	8	ADW32691	Adw32691 Bacterial
736	10	1.8	1468	4	ABB62991	Abb62991 Drosophil	809	9	1.6	556	4	ABB71805	Abb71805 Drosophil
737	10	1.8	1477	5	ABP69841	Abp69841 Human pol	810	9	1.6	558	4	ADW23595	Adw23595 Novel hum
738	10	1.8	1488	7	AAW37932	Aaw37932 Human CGP	811	9	1.6	607	7	ADW68340	Adw68340 Drosophil
739	10	1.8	1516	7	ABP69840	Abb69840 Human pol	812	9	1.6	607	7	ADW37889	Adw37889 D melano
740	10	1.8	1690	10	AEF41875	Aef41875 Human bet	813	9	1.6	628	2	ADW90059	Adw90059 Archaeabac
741	10	1.8	1695	9	AEW52608	Aew52608 Human bet	814	9	1.6	638	7	ADW90059	Adw90059 Archaeabac
742	10	1.8	1712	8	ABB60536	Abb60536 Drosophil	815	9	1.6	655	4	ABB65435	Abb65435 Drosophil
743	10	1.8	1905	8	ADW71535	Adw71535 Human CGD	816	9	1.6	662	4	ABB71837	Abb71837 Drosophil
744	10	1.8	2110	4	ABB58077	Abb58077 Drosophil	817	9	1.6	701	4	ABB57994	Abb57994 Drosophil
745	10	1.8	2586	4	ABB66878	Abb66878 Drosophil	818	9	1.6	710	8	ADW46583	Adw46583 Thermococ
746	10	1.8	4498	4	ABB58595	Abb58595 Drosophil	819	9	1.6	735	5	ABW73821	Abw73821 Candida a
747	10	1.8	8805	4	ABB67112	Abb67112 Drosophil	820	9	1.6	746	8	ADW30982	Adw30982 Human sec
748	9	1.6	10	4	AAW43213	Aaw43213 Mycoplaasm	821	9	1.6	774	4	ABB70674	Abb70674 Drosophil
749	9	1.6	10	4	AAW43223	Aaw43223 Mycoplaasm	822	9	1.6	777	4	ABB60154	Abb60154 Drosophil
750	9	1.6	10	4	AAW43229	Aaw43229 Mycoplaasm	823	9	1.6	798	7	ABW73728	Abw73728 Candida a
751	9	1.6	10	4	AAW43223	Aaw43223 Mycoplaasm	824	9	1.6	800	8	ABW58564	Abw58564 Human gen
752	9	1.6	10	4	AAW43249	Aaw43249 Mycoplaasm	825	9	1.6	801	4	ABB58990	Abb58990 Drosophil
753	9	1.6	51	3	AAW22340	Aaw22340 Arabidops	826	9	1.6	841	2	AAW34985	Aaw34985 Archaeabac

827	9	1.6	868	4	ABR65001	Abb65001 Drosophila	900	8	1.4	198	3	AA653458	Agg51458 Arabidops
828	9	1.6	941	5	ABP74093	Abp74093 Candida a	901	8	1.4	198	3	AA653227	Agg21227 Arabidops
829	9	1.6	1022	5	ADP99002	Adp99002 C. albica	902	8	1.4	202	3	AA653226	Agg21226 Arabidops
830	9	1.6	1077	8	ABR61013	Abbr61013 Drosophila	903	8	1.4	202	3	AA653457	Agg51457 Arabidops
831	9	1.6	1090	8	ADQ10187	Adq10187 Human pol	904	8	1.4	202	3	AA653367	Agg73367 Thale cre
832	9	1.6	1100	8	ADP99064	Adp99064 C. albica	905	8	1.4	204	3	AA651891	Agg11891 Arabidops
833	9	1.6	1158	4	ABR67681	Abbr67681 Drosophila	906	8	1.4	210	8	AD142173	Ad142173 Plant tra
834	9	1.6	1172	4	ABR64568	Abbr64568 Drosophila	907	8	1.4	210	8	AD002678	Ad002678 Thalecres
835	9	1.6	1404	2	AA838304	Aar33804 Sequence	908	8	1.4	210	8	AD062812	Ad062812 Transcrip
836	9	1.6	1404	3	AA596600	Aay59600 Drosophila	909	8	1.4	211	4	ABU71740	Abu71740 Drosophila
837	9	1.6	1404	3	ABR61998	Abbr61998 Drosophila	910	8	1.4	211	4	ADU02385	Adu02385 Novel hum
838	9	1.6	1404	5	ABR07827	Abbr07827 Drosophila	911	8	1.4	219	6	ABU33737	Abu33737 Protein e
839	9	1.6	1404	8	ADL71360	Adl71360 Seriate p	912	8	1.4	220	4	ABR69176	Abbr69176 Drosophila
840	9	1.6	1542	6	ABR71456	Abbr71456 Drosophila	913	8	1.4	222	10	AEF11608	Aef11608 Soybean m
841	9	1.6	1652	6	ADA15715	Ada15715 C. elegans	914	8	1.4	225	8	ADP30589	Adp30589 Human sec
842	9	1.6	1652	10	AEF43012	Aef43012 Nematode	915	8	1.4	225	8	ADP30589	Adp30589 Human sec
843	9	1.6	1795	4	ABR69806	Abbr69806 Drosophila	916	8	1.4	239	8	ADP30589	Adp30589 Human sec
844	9	1.6	1917	6	ADA15719	Ada15719 C. elegans	917	8	1.4	241	7	ABR69806	Abbr69806 Rice abio
845	9	1.6	1917	10	AEF43016	Aef43016 Nematode	918	8	1.4	254	3	AA653780	Agg53780 Arabidops
846	9	1.6	2424	4	ABR69924	Abbr69924 Nematode	919	8	1.4	257	3	AA653780	Agg53780 Arabidops
847	9	1.6	2508	10	AEF43018	Aef43018 Nematode	920	8	1.4	272	2	AA653780	Agg53780 Arabidops
848	9	1.6	2508	10	AEF43018	Aef43018 Nematode	921	8	1.4	272	2	AA653780	Agg53780 Arabidops
849	9	1.6	2544	6	ADA15717	Ada15717 C. elegans	922	8	1.4	273	7	ABR69924	Abbr69924 Rice abio
850	9	1.6	2544	10	AEF43014	Aef43014 Nematode	923	8	1.4	277	2	AA653193	Agg53193 Arabidops
851	9	1.6	2601	6	ADA15723	Ada15723 C. elegans	924	8	1.4	287	2	AA653193	Agg53193 Arabidops
852	9	1.6	2601	10	AEF43020	Aef43020 Nematode	925	8	1.4	290	9	ADN18269	Adn18269 Eucalyptu
853	9	1.6	2638	8	ADP30981	Adp30981 Human sec	926	8	1.4	296	4	ABR69806	Abbr69806 Rice abio
854	9	1.6	3672	8	ADN23493	Adn23493 Bacterial	927	8	1.4	300	2	AA653780	Agg53780 Arabidops
855	9	1.4	3672	8	ADN23493	Adn23493 Bacterial	928	8	1.4	323	7	ADP30589	Adp30589 Human sec
856	9	1.4	10	10	AEF53946	Aef53946 Size stan	929	8	1.4	323	7	ADP30589	Adp30589 Human sec
857	9	1.4	10	10	AEF53946	Aef53946 Size stan	930	8	1.4	323	7	ADP30589	Adp30589 Human sec
858	9	1.4	10	10	AEF53946	Aef53946 Size stan	931	8	1.4	323	7	ADP30589	Adp30589 Human sec
859	9	1.4	10	10	AEF53946	Aef53946 Size stan	932	8	1.4	324	2	AA653780	Agg53780 Arabidops
860	9	1.4	10	10	AEF53946	Aef53946 Size stan	933	8	1.4	327	2	AA653780	Agg53780 Arabidops
861	9	1.4	10	10	AEF53946	Aef53946 Size stan	934	8	1.4	331	3	AA653780	Agg53780 Arabidops
862	9	1.4	10	10	AEF53946	Aef53946 Size stan	935	8	1.4	333	3	AA653780	Agg53780 Arabidops
863	9	1.4	10	10	AEF53946	Aef53946 Size stan	936	8	1.4	333	3	AA653780	Agg53780 Arabidops
864	9	1.4	10	10	AEF53946	Aef53946 Size stan	937	8	1.4	333	3	AA653780	Agg53780 Arabidops
865	9	1.4	10	10	AEF53946	Aef53946 Size stan	938	8	1.4	333	3	AA653780	Agg53780 Arabidops
866	9	1.4	10	10	AEF53946	Aef53946 Size stan	939	8	1.4	333	3	AA653780	Agg53780 Arabidops
867	9	1.4	10	10	AEF53946	Aef53946 Size stan	940	8	1.4	333	3	AA653780	Agg53780 Arabidops
868	9	1.4	10	10	AEF53946	Aef53946 Size stan	941	8	1.4	333	3	AA653780	Agg53780 Arabidops
869	9	1.4	10	10	AEF53946	Aef53946 Size stan	942	8	1.4	333	3	AA653780	Agg53780 Arabidops
870	9	1.4	10	10	AEF53946	Aef53946 Size stan	943	8	1.4	333	3	AA653780	Agg53780 Arabidops
871	9	1.4	10	10	AEF53946	Aef53946 Size stan	944	8	1.4	333	3	AA653780	Agg53780 Arabidops
872	9	1.4	10	10	AEF53946	Aef53946 Size stan	945	8	1.4	333	3	AA653780	Agg53780 Arabidops
873	9	1.4	10	10	AEF53946	Aef53946 Size stan	946	8	1.4	333	3	AA653780	Agg53780 Arabidops
874	9	1.4	10	10	AEF53946	Aef53946 Size stan	947	8	1.4	333	3	AA653780	Agg53780 Arabidops
875	9	1.4	10	10	AEF53946	Aef53946 Size stan	948	8	1.4	333	3	AA653780	Agg53780 Arabidops
876	9	1.4	10	10	AEF53946	Aef53946 Size stan	949	8	1.4	333	3	AA653780	Agg53780 Arabidops
877	9	1.4	10	10	AEF53946	Aef53946 Size stan	950	8	1.4	333	3	AA653780	Agg53780 Arabidops
878	9	1.4	10	10	AEF53946	Aef53946 Size stan	951	8	1.4	333	3	AA653780	Agg53780 Arabidops
879	9	1.4	10	10	AEF53946	Aef53946 Size stan	952	8	1.4	333	3	AA653780	Agg53780 Arabidops
880	9	1.4	10	10	AEF53946	Aef53946 Size stan	953	8	1.4	333	3	AA653780	Agg53780 Arabidops
881	9	1.4	10	10	AEF53946	Aef53946 Size stan	954	8	1.4	333	3	AA653780	Agg53780 Arabidops
882	9	1.4	10	10	AEF53946	Aef53946 Size stan	955	8	1.4	333	3	AA653780	Agg53780 Arabidops
883	9	1.4	10	10	AEF53946	Aef53946 Size stan	956	8	1.4	333	3	AA653780	Agg53780 Arabidops
884	9	1.4	10	10	AEF53946	Aef53946 Size stan	957	8	1.4	333	3	AA653780	Agg53780 Arabidops
885	9	1.4	10	10	AEF53946	Aef53946 Size stan	958	8	1.4	333	3	AA653780	Agg53780 Arabidops
886	9	1.4	10	10	AEF53946	Aef53946 Size stan	959	8	1.4	333	3	AA653780	Agg53780 Arabidops
887	9	1.4	10	10	AEF53946	Aef53946 Size stan	960	8	1.4	333	3	AA653780	Agg53780 Arabidops
888	9	1.4	10	10	AEF53946	Aef53946 Size stan	961	8	1.4	333	3	AA653780	Agg53780 Arabidops
889	9	1.4	10	10	AEF53946	Aef53946 Size stan	962	8	1.4	333	3	AA653780	Agg53780 Arabidops
890	9	1.4	10	10	AEF53946	Aef53946 Size stan	963	8	1.4	333	3	AA653780	Agg53780 Arabidops
891	9	1.4	10	10	AEF53946	Aef53946 Size stan	964	8	1.4	333	3	AA653780	Agg53780 Arabidops
892	9	1.4	10	10	AEF53946	Aef53946 Size stan	965	8	1.4	333	3	AA653780	Agg53780 Arabidops
893	9	1.4	10	10	AEF53946	Aef53946 Size stan	966	8	1.4	333	3	AA653780	Agg53780 Arabidops
894	9	1.4	10	10	AEF53946	Aef53946 Size stan	967	8	1.4	333	3	AA653780	Agg53780 Arabidops
895	9	1.4	10	10	AEF53946	Aef53946 Size stan	968	8	1.4	333	3	AA653780	Agg53780 Arabidops
896	9	1.4	10	10	AEF53946	Aef53946 Size stan	969	8	1.4	333	3	AA653780	Agg53780 Arabidops
897	9	1.4	10	10	AEF53946	Aef53946 Size stan	970	8	1.4	333	3	AA653780	Agg53780 Arabidops
898	9	1.4	10	10	AEF53946	Aef53946 Size stan	971	8	1.4	333	3	AA653780	Agg53780 Arabidops
899	9	1.4	10	10	AEF53946	Aef53946 Size stan	972	8	1.4	333	3	AA653780	Agg53780 Arabidops

973	8	1.4	543	5	ABP73740	Abp73740	Candida a
974	8	1.4	543	7	ABM88929	Abm88929	Rice abio
975	8	1.4	551	8	ADK68546	Adk68546	Plant ful
976	8	1.4	557	8	ADS44712	AdS44712	Bacterial
977	8	1.4	558	6	ADA17078	Ada17078	Nuclear f
978	8	1.4	558	6	ADA17079	Ada17079	Nuclear f
979	8	1.4	562	4	ABBS9914	Abbs9914	Drosophi1
980	8	1.4	587	3	AAB85914	Aab85914	Arabidops
981	8	1.4	587	4	AAE01892	Aae01892	Arabidops
982	8	1.4	587	4	AAE02560	Aae02560	A. thalia
983	8	1.4	587	8	AD001803	Ad001803	Thalictres
984	8	1.4	587	9	AEA27333	Aea27333	Streps to
985	8	1.4	592	4	AA882297	Aab82297	Arabidops
986	8	1.4	592	4	AA882296	Aab82296	Arabidops
987	8	1.4	592	5	ABP99405	Abp99405	Arabidops
988	8	1.4	592	8	ADM98799	Adm98799	HMG-COA r
989	8	1.4	592	8	ADM98789	Adm98789	HMG-COA r
990	8	1.4	592	8	ADM98837	Adm98837	HMG-COA r
991	8	1.4	592	8	ADM98827	Adm98827	HMG-COA r
992	8	1.4	592	9	ADY52924	Ady52924	Thale cre
993	8	1.4	592	9	ADY52374	Ady52374	Novel ket
994	8	1.4	592	9	ADY51371	Ady51371	Arabidops
995	8	1.4	595	4	ABB71737	Abb71737	Drosophi1
996	8	1.4	600	4	ABB63003	Abb63003	Drosophi1
997	8	1.4	603	4	ABB65941	Abb65941	Drosophi1
998	8	1.4	620	5	ABJ05379	Abj05379	Chimeric
999	8	1.4	620	7	ADF49209	Adf49209	Ecdysone
1000	8	1.4	623	2	AAR66875	Aar66875	Cancer su

ALIGNMENTS

```
RESULT 1
ADC00799
ID   ADC00799 standard; protein; 558 AA.
XX
AC   ADC00799;
XX
DT   04-DEC-2003 (first entry)
DE   Enterohaemorrhagic E. coli O157:H7-specific protein SEQ ID NO: 844.
XX
KW   enterohaemorrhagic; anti-bacterial.
XX
OS   Escherichia coli; O157:H7.
XX
PN   JP200235074-A.
XX
PD   10-DEC-2002.
XX
PF   24-JAN-2002; 2002JP-00015959.
XX
PR   24-JAN-2001; 2001JP-00112010.
XX
PA   (UYTS-) UNIV TSUKUBA.
XX
DR   WPI; 2003-451640/43.
XX
PT   Enterohaemorrhagic Escherichia coli O157:H7-specific nucleic acid molecule
XX   and a polypeptide and its use, a polypeptide, a vector and a host cell.
XX
PS   Claim 3; SEQ ID NO 844; 2067pp; Japanese.
XX
CC   The invention relates to a novel enterohaemorrhagic Escherichia coli
CC   O157:H7-specific nucleic acid molecule. A polynucleotide of the invention
CC   has anti-bacterial activity. The polypeptide can be used in detection
CC   and/or treatment of O157:H7 infection. The nucleotide sequence of the
CC   genome of Enterohaemorrhagic E coli O157:H7 was determined. The present
CC   sequence represents an E. coli O157:H7-specific polypeptide of the
XX   invention.
SQ   Sequence 558 AA;
```

Query Match	100.0%	Score 558	DB 7	Length 558	
Beet Local Similarity	100.0%	Pred. No. 0			
Matches 558; Conservative	0	Mismatches	0	Indels	0
Gaps	0				
QY	1	MEIGNLGNPNVNNISIPAPPLPSQTDGAGRGQILNSTGPLSGHALFTPVNNSMADSGD	60		
DB	1	MPIGNLGNPNVNNISIPAPPLPSQTDGAGRGQILNSTGPLSGHALFTPVNNSMADSGD	60		
QY	61	NRASVPGLPVNPMPLAASEITLNDGFEVLHDHGGLDNLNQIGSSVFRVETQEDGKHIA	120		
DB	61	NRASVPGLPVNPMPLAASEITLNDGFEVLHDHGGLDNLNQIGSSVFRVETQEDGKHIA	120		
QY	121	VGORNGVETSVVLSDOEFARLQSIDPEGDKFVFETGRGAGHAMVTAASDITTEARQIL	180		
DB	121	VGORNGVETSVVLSDOEFARLQSIDPEGDKFVFETGRGAGHAMVTAASDITTEARQIL	180		
QY	181	ELLEPKGTGESKGAESKGVGELRESNSGAENTTETQSTSTSSLRSPDKMLALGTVAAT	240		
DB	181	ELLEPKGTGESKGAESKGVGELRESNSGAENTTETQSTSTSSLRSPDKMLALGTVAAT	240		
QY	241	GLIGLAATGIYQALALTPEDPSPTTDPDAAASATETATROQLTEAFONPNQKVNIDE	300		
DB	241	GLIGLAATGIYQALALTPEDPSPTTDPDAAASATETATROQLTEAFONPNQKVNIDE	300		
QY	301	LGNAIPSGVLKDDVVANIEQAKAAGEBAKQOAIENNAQAKKYDEQQAQKEBELKVS	360		
DB	301	LGNAIPSGVLKDDVVANIEQAKAAGEBAKQOAIENNAQAKKYDEQQAQKEBELKVS	360		
QY	361	AGYGLSGALLGGGIGVAVTALHKNQPVBEQTTTTTTTTTTTTSARVYENKPAANTPAQ	420		
DB	361	AGYGLSGALLGGGIGVAVTALHKNQPVBEQTTTTTTTTTTTTSARVYENKPAANTPAQ	420		
QY	421	NVDTPGSEDTMESRRSSMASTSTFTDSSIGTVONPYADVKTSLHDSQVPTSNSTSVQ	480		
DB	421	NVDTPGSEDTMESRRSSMASTSTFTDSSIGTVONPYADVKTSLHDSQVPTSNSTSVQ	480		
QY	481	NMGNTDSVYVSTIQHPPRDTTNGARLLGNPSAGIQSTYARLALSGLRHDMGGLTGSGN	540		
DB	481	NMGNTDSVYVSTIQHPPRDTTNGARLLGNPSAGIQSTYARLALSGLRHDMGGLTGSGN	540		
QY	541	SAVNTSNPPAPGSHRFV	558		
DB	541	SAVNTSNPPAPGSHRFV	558		
RESULT 2					
AE91310	standard; protein; 558 AA.				
XX					
AC	AE91310;				
XX					
DT	20-OCT-2005 (first entry)				
XX					
DE	Microbial pathogen adhesin protein sequence, SEQ ID NO:20.				
XX					
KW	algorithm; adhesin; pharmaceutical; vaccine; drug screening;				
KW	bordelella pertussis infection; antibacterial; pneumonia;				
KW	antiinflammatory; respiratory-gen.; gastric ulcer; antiulcer;				
KW	gastrointestinal-gen.; urinary tract infection; antimicrobial; uropathic.				
OS	Escherichia coli.				
XX					
PN	WO2005076010-A2.				
XX					
PD	18-AUG-2005.				
XX					
PF	07-FEB-2005; 2005WO-IN000037.				
XX					
PR	06-FEB-2004; 2004IN-DE000173.				
XX					
PR	20-JUL-2004; 2004US-0589227P.				
XX					
PA	(COUL) COUNCIL SCI & IND RES SOUTH AFRICA.				

XX Sachdeva G, Kumar K, Jain P, Brahmachari SK, Ramachandran S;
 PI MPI; 2005-597835/61.
 XX
 DR Computational method for identifying adhesin and adhesin like molecules,
 PT comprises computing sequence-based attributes of protein sequences using
 PT neural network software and training an artificial neural network.
 XX
 PS Claim 16; SEQ ID NO 20; 402pp; English.
 XX
 CC The present invention relates to a computational method (M1) for
 CC identifying adhesin and adhesin-like proteins, by computing the sequence-
 CC based attributes of protein sequences using five attribute modules of a
 CC neural network software, training an artificial neural network (ANN) for
 CC each of the computed five attributes, and identifying the adhesin and
 CC adhesin-like proteins having probability of being an adhesin (Pad) as
 CC equal or greater than 0.51. Also claimed is a set of 274 annotated genes
 CC encoding adhesin and adhesin-like proteins, having 274 fully defined 162-
 CC 1151 base pair (SEQ ID NO: 385-658) sequences; a set of 105 hypothetical
 CC genes encoding adhesin and adhesin-like proteins, having 105 fully
 CC defined 306-1576 base pairs (SEQ ID NO: 659-763); a set of 279 annotated
 CC adhesin and adhesin-like proteins, having 279 fully defined 53-3716 base
 CC pair (SEQ ID NO: 1-279); a set of 105 hypothetical adhesin and adhesin-
 CC like proteins, having 105 fully defined 106-5291 base pair (SEQ ID NO:
 CC 280-384) sequences; and a fully connected multilayer feed forward ANN (1)
 CC based on (M1). (M1) is useful for identifying adhesin and adhesin-like
 CC proteins, of therapeutic potential, and identifying and short-listing
 CC proteins for further testing in development of new vaccine formulations
 CC to eliminate diseases caused by various pathogenic organisms. (M1) is
 CC useful for identifying putative adhesins that are important in drug
 CC discovery and preventing therapeutics for whooping cough, pneumonia,
 CC gastric ulcer and urinary tract infections. (M1) identifies adhesins from
 CC distantly related organisms, and from bacteria belonging to a wide
 CC phylogenetic spectrum. (M1) is capable of predicting adhesive nature of
 CC unique proteins. The present sequence is a microbial pathogen adhesin
 CC protein sequence.
 CC
 XX Sequence 558 AA;
 SO
 Query Match 100.0%; Score 558; DB 9; Length 558;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 558; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MPICGLGNHPNNVNSIPPAPPLPSQTDGAGRGQLNSTGPGSRALFPVNNMSADSD 60
 DB 1 MPICGLGNHPNNVNSIPPAPPLPSQTDGAGRGQLNSTGPGSRALFPVNNMSADSD 60
 QY 61 NRASVDPGLPVNPMRLAASEITLNDGFEVLHDHGPDLTNRQIGSSVFVEVOEGCKHIA 120
 DB 61 NRASVDPGLPVNPMRLAASEITLNDGFEVLHDHGPDLTNRQIGSSVFVEVOEGCKHIA 120
 QY 121 VQQRNGVETSVVLSDOEVARLQSIDPEGKDFEFTGSRGAGAHAMTVASDITTEARORTL 180
 DB 121 VQQRNGVETSVVLSDOEVARLQSIDPEGKDFEFTGSRGAGAHAMTVASDITTEARORTL 180
 QY 121 VQQRNGVETSVVLSDOEVARLQSIDPEGKDFEFTGSRGAGAHAMTVASDITTEARORTL 180
 DB 121 VQQRNGVETSVVLSDOEVARLQSIDPEGKDFEFTGSRGAGAHAMTVASDITTEARORTL 180
 QY 181 ELLEKRGKGESKGAESKGVGLRBSNSGAENTTETOTSTSSLSRDPKMLALGVAT 240
 DB 181 ELLEKRGKGESKGAESKGVGLRBSNSGAENTTETOTSTSSLSRDPKMLALGVAT 240
 QY 181 ELLEKRGKGESKGAESKGVGLRBSNSGAENTTETOTSTSSLSRDPKMLALGVAT 240
 DB 181 ELLEKRGKGESKGAESKGVGLRBSNSGAENTTETOTSTSSLSRDPKMLALGVAT 240
 QY 241 GLIGLAATGIYVALLTPEPDSPTTTPDAASATETANRDLTTEAFPNPNQVNI 300
 DB 241 GLIGLAATGIYVALLTPEPDSPTTTPDAASATETANRDLTTEAFPNPNQVNI 300
 QY 241 GLIGLAATGIYVALLTPEPDSPTTTPDAASATETANRDLTTEAFPNPNQVNI 300
 DB 241 GLIGLAATGIYVALLTPEPDSPTTTPDAASATETANRDLTTEAFPNPNQVNI 300
 QY 301 LGNALPSGVLKDDVVANIEEQAKAGBEAKQQAIEENNAQAKYEQQAQKQBEELKVS 360
 DB 301 LGNALPSGVLKDDVVANIEEQAKAGBEAKQQAIEENNAQAKYEQQAQKQBEELKVS 360
 QY 301 LGNALPSGVLKDDVVANIEEQAKAGBEAKQQAIEENNAQAKYEQQAQKQBEELKVS 360
 DB 301 LGNALPSGVLKDDVVANIEEQAKAGBEAKQQAIEENNAQAKYEQQAQKQBEELKVS 360
 QY 361 AGYGSGLAILLGGGIGVAVTALALHRKNQPEVOTTTTTTTTTTSARVYENKPAANTPA 420
 DB 361 AGYGSGLAILLGGGIGVAVTALALHRKNQPEVOTTTTTTTTTTSARVYENKPAANTPA 420
 QY 421 NVDTGSEDTWBSRRSSMASTSTFPDSSIGTVQNPADVATSLHDGQVFPISNNTSVQ 480

DB 421 NVDTGSEDTWBSRRSSMASTSTFPDSSIGTVQNPADVATSLHDGQVFPISNNTSVQ 480
 QY 481 NMGNTDSVYVSTIIOHPERDITDNGARLLGNPSAGISTYARLALSGGLRHHMGGLTGGSN 540
 DB 481 NMGNTDSVYVSTIIOHPERDITDNGARLLGNPSAGISTYARLALSGGLRHHMGGLTGGSN 540
 QY 541 SAVNTSNPNPAPGSHRFV 558
 DB 541 SAVNTSNPNPAPGSHRFV 558
 RESULT 3
 ABE86220
 ID ABE86220 standard; protein; 558 AA.
 XX
 AC ABE86220;
 XX
 DT 23-FEB-2006 (first entry)
 XX
 DE Escherichia coli translocated intimin receptor (Tir) protein.
 XX
 KW Diagnosis; therapeutic; screening; escherichia coli infection;
 KW antibacterial; infection; translocated intimin receptor.
 XX
 OS Escherichia coli.
 XX
 PN US2005287569-A1.
 XX
 PD 29-DEC-2005.
 XX
 PF 20-MAY-2005; 2005US-00134563.
 XX
 PR 20-MAY-2004; 2004US-0573600P.
 XX
 PA (LEONG/) LEONG J M.
 PA (CAMP/) CAMPBELLONE K G.
 XX
 PI Leong JM, Campellone KG,
 XX
 DR MPI; 2006-065745/07.
 XX
 N-PSDB; ABE86219.
 PT Novel purified polypeptide having six residues of EspF-U, and binding to
 PT neuronal Wiskott-Aldrich syndrome protein polypeptide, useful for
 PT identifying EspF-U activity modulating compound.
 PS Disclosure; SEQ ID NO 12; 62pp; English.
 XX
 CC The present invention relates to novel EspFu polypeptides and their
 CC corresponding polynucleotides. The EspFu polypeptides are EspF-like
 CC polypeptides encoded by genes of the cryptic prophage CP-933U of
 CC enterohemorrhagic Escherichia coli (EHEC) or enteropathogenic Escherichia
 CC coli (EPEC) that binds to a neuronal Wiskott-Aldrich syndrome protein (N-
 CC WASP) polypeptide or restores the actin pedestal formation activity of
 CC enteropathogenic E. coli (EPEC) strain K12. The invention further
 CC relates to a method of identifying a candidate compound capable of
 CC binding to and/or modulating the activity of EspFu and compounds that
 CC inhibits protein-protein interactions between EspFu and EspFu-interacting
 CC proteins such as N-WASP. Transducer of Cdc42-dependent actin assembly-1
 CC (Toca-1) and p21-activated kinase 1 (Pak1). EspFu polynucleotides are
 CC useful for diagnosing or detecting EHEC infection. EspFu antibody is
 CC useful for treating EHEC infection. The present sequence is the
 CC enterohemorrhagic Escherichia coli translocated intimin receptor (Tir)
 CC protein. This sequence is critical for the formation of actin pedestals
 CC in EHEC.
 CC
 XX Sequence 558 AA;
 SO
 Query Match 100.0%; Score 558; DB 10; Length 558;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 558; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
OY 1 MPIGNLGNPNVNNISIPAPPLPSQTDGAGRGQLINSTGPLGSALFTPVNSMADSGD 60
DB 1 MPIGNLGNPNVNNISIPAPPLPSQTDGAGRGQLINSTGPLGSALFTPVNSMADSGD 60
OY 61 NRASDVPGLPVNPMLAASEITLNDGFEVLHDHGPLDITLNROIGSSVFRVETOEDGKHIA 120
DB 61 NRASDVPGLPVNPMLAASEITLNDGFEVLHDHGPLDITLNROIGSSVFRVETOEDGKHIA 120
OY 121 VGORNGVETSVLSDQEVARLOSIDPEKDKFVFTGGRGAGHAMVTVASDITEARORIL 180
DB 121 VGORNGVETSVLSDQEVARLOSIDPEKDKFVFTGGRGAGHAMVTVASDITEARORIL 180
OY 181 ELLEPKGTGESKGAESKGVGELRESNSGAENTTETOTSTSTSSLRSDPKMLALGTAT 240
DB 181 ELLEPKGTGESKGAESKGVGELRESNSGAENTTETOTSTSTSSLRSDPKMLALGTAT 240
OY 241 GLIGLAATGIVQALALTPEDPSPTTTDPDAASATEFATRDQLTKEAFQNPDPNOKVNIDE 300
DB 241 GLIGLAATGIVQALALTPEDPSPTTTDPDAASATEFATRDQLTKEAFQNPDPNOKVNIDE 300
OY 301 LGNALPSGVLKDDVVANIEEQAKAAGEAKQOAIENNAQAQKXDEQAQKQEBLKVSNG 360
DB 301 LGNALPSGVLKDDVVANIEEQAKAAGEAKQOAIENNAQAQKXDEQAQKQEBLKVSNG 360
OY 361 AGYGLSGALLIGGGIGVAVTAAHHRKNQVVEOTTTTTTTTTTSARTVENKPAANTPAQG 420
DB 361 AGYGLSGALLIGGGIGVAVTAAHHRKNQVVEOTTTTTTTTTTSARTVENKPAANTPAQG 420
OY 421 NVDTPGSEDTMESRRSSMASTSTFTFTSSIGTVONPVADVTSLLHDSQVPTSNSNTSVQ 480
DB 421 NVDTPGSEDTMESRRSSMASTSTFTFTSSIGTVONPVADVTSLLHDSQVPTSNSNTSVQ 480
OY 481 NMGNQDSVVYSTIQQHPPRDTTNGARLLGNPSAGIOSTYARLALSGGLSHDMGGLTGSN 540
DB 481 NMGNQDSVVYSTIQQHPPRDTTNGARLLGNPSAGIOSTYARLALSGGLSHDMGGLTGSN 540
OY 541 SAVNTSNNPPAPGSHRFV 558
DB 541 SAVNTSNNPPAPGSHRFV 558

RESULT 4
ID AAY06221 standard; protein; 559 AA.
XX
AC AAY06221;
XX
DT 16-AUG-1999 (first entry)
XX
DE EHEC E. coli translocated intimin receptor (Tir).
XX
KM Tir; translocated intimin receptor; Hp90; enterohaemorrhagic; EHEC;
XX infection; diagnosis; vaccine.
XX
OS Escherichia coli.
XX
XX
FH Key Location/Qualifiers
FT Misc-difference 453 /note="encoded by codon of 1 apparent nucleotide,
XX causing frameshift in the DNA sequence"
XX
PN WO924576-A1.
XX
PD 20-MAY-1999.
XX
PF 10-NOV-1998; 98WO-CA001042.
XX
PR 12-NOV-1997; 97US-0065130P.
XX
XX (UYBR-) UNIV BRITISH COLUMBIA.
XX
PA
XX
PI Finlay BB, Kenny B, Devlinney R, Stein M;
XX
```

```
DR WPI; 1999-337712/28.
DR N-PSDB; AAX58859.
XX
PT New translocated intimin receptor useful for treating infection by
PT enteropathogenic or enterohaemorrhagic Escherichia coli.
XX
PS Claim 7; Page 55-58; 91pp; English.
XX
CC The present sequence represents Tir, a novel translocated intimin
CC receptor (formerly termed Hp90) from an enterohaemorrhagic Escherichia
CC coli (EHEC) strain. The sequence was deduced from an isolated tir
CC polynucleotide (see AAX58859). Tir proteins are secreted by attaching and
CC effecting pathogens such as EHEC and EPEC (see AAY06220) E. coli. The
CC bacterial pathogens insert their own receptors into mammalian cell
CC surfaces, to which the pathogen then adheres to trigger additional host
CC signaling events and actin nucleation. Diagnosis of disease caused by
CC pathogenic E. coli can be performed by use of antibodies that bind to Tir
CC to detect the protein or the use of nucleic acid probes for detection of
CC nucleic acids encoding Tir polypeptide. Isolated Tir nucleic acids, Tir
CC peptides, a recombinant method for producing recombinant Tir, antibodies
CC which bind to Tir, and a kit for the detection of Tir-producing E. coli
CC are provided. A method of immunising a host with Tir to induce a
CC protective immune response is also provided. In addition, Tir fusion
CC proteins can be used in attenuated E. coli to induce a cell-mediated
CC immune response to other polypeptides, e.g. antigens. A method for
CC screening for compounds which interfere with the binding of bacterial
CC pathogens to their receptors is further provided
XX
SQ Sequence 559 AA;
```

```
Query Match 81.0%; Score 452; DB 2; Length 559;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 452; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 MPIGNLGNPNVNNISIPAPPLPSQTDGAGRGQLINSTGPLGSALFTPVNSMADSGD 60
DB 1 MPIGNLGNPNVNNISIPAPPLPSQTDGAGRGQLINSTGPLGSALFTPVNSMADSGD 60
OY 61 NRASDVPGLPVNPMLAASEITLNDGFEVLHDHGPLDITLNROIGSSVFRVETOEDGKHIA 120
DB 61 NRASDVPGLPVNPMLAASEITLNDGFEVLHDHGPLDITLNROIGSSVFRVETOEDGKHIA 120
OY 121 VGORNGVETSVLSDQEVARLOSIDPEKDKFVFTGGRGAGHAMVTVASDITEARORIL 180
DB 121 VGORNGVETSVLSDQEVARLOSIDPEKDKFVFTGGRGAGHAMVTVASDITEARORIL 180
OY 181 ELLEPKGTGESKGAESKGVGELRESNSGAENTTETOTSTSTSSLRSDPKMLALGTAT 240
DB 181 ELLEPKGTGESKGAESKGVGELRESNSGAENTTETOTSTSTSSLRSDPKMLALGTAT 240
OY 241 GLIGLAATGIVQALALTPEDPSPTTTDPDAASATEFATRDQLTKEAFQNPDPNOKVNIDE 300
DB 241 GLIGLAATGIVQALALTPEDPSPTTTDPDAASATEFATRDQLTKEAFQNPDPNOKVNIDE 300
OY 301 LGNALPSGVLKDDVVANIEEQAKAAGEAKQOAIENNAQAQKXDEQAQKQEBLKVSNG 360
DB 301 LGNALPSGVLKDDVVANIEEQAKAAGEAKQOAIENNAQAQKXDEQAQKQEBLKVSNG 360
OY 361 AGYGLSGALLIGGGIGVAVTAAHHRKNQVVEOTTTTTTTTTTSARTVENKPAANTPAQG 420
DB 361 AGYGLSGALLIGGGIGVAVTAAHHRKNQVVEOTTTTTTTTTTSARTVENKPAANTPAQG 420
OY 421 NVDTPGSEDTMESRRSSMASTSTFTFTSSIG 452
DB 421 NVDTPGSEDTMESRRSSMASTSTFTFTSSIG 452
```

```
RESULT 5
ID AAY06213 standard; peptide; 30 AA.
XX
AC AAY06213;
XX
```

```
DT 16-AUG-1999 (first entry)
XX
XX EPEC E. coli translocated intimin receptor N-terminal peptide.
DE
XX T1r; translocated intimin receptor; Hp90; enteropathogenic; EPEC;
KM infection; diagnosis; vaccine.
XX
XX Escherichia coli.
OS
XX MO9924576-A1.
XX
XX 20-MAY-1999.
XX
XX 10-NOV-1998; 98MO-CA001042.
XX
XX 12-NOV-1997; 97US-0065130P.
XX
XX (UYBR-) UNIV BRITISH COLUMBIA.
XX
XX Finlay BB, Kenny B, Devinney R, Stein M;
XX
XX MPI; 1999-337712/28.
XX
XX New translocated intimin receptor useful for treating infection by
PT enteropathogenic or enterohemorrhagic Escherichia coli.
XX
XX Example 1, Page 37; 91pp; English.
XX
XX The present sequence represents the N-terminal sequence of T1r (see also
CC AAY06220), a novel translocated intimin receptor from an enteropathogenic
CC Escherichia coli (EPEC) strain. The 78 kDa EPEC protein is secreted by
CC the bacterial pathogen. Diagnosis of disease caused by pathogenic E. coli
CC can be performed by use of antibodies that bind to T1r to detect the
CC protein, or the use of nucleic acid probes for detection of nucleic acids
CC encoding T1r. A kit for the detection of T1r-producing E. coli is
CC provided. Also provided are a method of immunising a host with T1r to
CC induce a protective immune response, and a method for screening for
CC compounds which interfere with the binding of bacterial pathogens to
CC their receptors
XX
XX Sequence 30 AA:
SQ
Query Match 2.5%; Score 14; DB 2; Length 30;
Best Local Similarity 100.0%; Pred. No. 1.9e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 16 IPPAPPLPSQTDGA 29
DB 15 IPPAPPLPSQTDGA 28

RESULT 6
AAY06220 standard; protein; 549 AA.
XX
XX AAY06220;
XX
XX 16-AUG-1999 (first entry)
XX
XX EPEC E. coli translocated intimin receptor (T1r).
DE
XX T1r; translocated intimin receptor; Hp90; enteropathogenic; EPEC;
KM infection; diagnosis; vaccine.
XX
XX Escherichia coli.
OS
XX
XX Key Location/Qualifiers
XX Misc-difference 180
XX FT /note= "encoded by AAA"
XX FT 234..253
XX FT /note= "putative transmembrane domain"
XX FT Misc-difference 314
XX FT /note= "given as Xaa in the specification; Lys is deduced
```

```
FT FT from the DNA sequence"
FT Domain 364..386
FT FT /note= "putative transmembrane domain"
XX
XX MO9924576-A1.
XX
XX 20-MAY-1999.
XX
XX 10-NOV-1998; 98MO-CA001042.
XX
XX 12-NOV-1997; 97US-0065130P.
XX
XX (UYBR-) UNIV BRITISH COLUMBIA.
XX
XX Finlay BB, Kenny B, Devinney R, Stein M;
XX
XX MPI; 1999-337712/28.
XX
XX N-PSDB; AAX58658.
XX
XX New translocated intimin receptor useful for treating infection by
PT enteropathogenic or enterohemorrhagic Escherichia coli.
XX
XX Claim 6; Page 55-58; 91pp; English.
XX
XX The present sequence represents T1r, a novel translocated intimin
CC receptor (formerly termed Hp90) from an enteropathogenic Escherichia coli
CC (EPEC) strain. The sequence was deduced from an isolated t1r
CC polynucleotide (see AAX58858). T1r proteins are secreted by attaching and
CC effecting pathogens such as EPEC and EHEC (see AAY06221) E. coli. The
CC bacterial pathogens insert their own receptors into mammalian cell
CC surfaces, to which the pathogen then adheres to trigger additional host
CC signaling events and actin nucleation. Diagnosis of disease caused by
CC pathogenic E. coli can be performed by use of antibodies that bind to T1r
CC to detect the protein or the use of nucleic acid probes for detection of
CC nucleic acids encoding T1r polypeptide. Isolated T1r nucleic acids, T1r
CC peptides, a recombinant method for producing recombinant T1r, antibodies
CC which bind to T1r, and a kit for the detection of T1r-producing E. coli
CC are provided. A method of immunising a host with T1r to induce a
CC protective immune response is also provided. In addition, T1r fusion
CC proteins can be used in attenuated E. coli to induce a cell-mediated
CC immune response to other polypeptides, e.g. antigens. A method for
CC screening for compounds which interfere with the binding of bacterial
CC pathogens to their receptors is further provided
XX
XX Sequence 549 AA:
SQ
Query Match 2.5%; Score 14; DB 2; Length 549;
Best Local Similarity 100.0%; Pred. No. 0.00028;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 16 IPPAPPLPSQTDGA 29
DB 16 IPPAPPLPSQTDGA 29

RESULT 7
AEE26469
ID AEE26469 standard; protein; 162 AA.
XX
XX AEE26469;
XX
XX 09-FEB-2006 (first entry)
XX
XX ORF0826 immunogen related sequence #1.
XX
XX Vaccine: Gene-therapy; Antibacterial; immunogen; ORF0826; Ssa;
XX Staphylococcus aureus; infection.
XX
XX Staphylococcus aureus; methicillin resistant strain.
XX
XX Synthetic.
XX
XX MO2005115113-A2.
XX
```

PD 08-DEC-2005.
XX
PF 20-MAY-2005; 2005WO-US017835.
XX
PR 25-MAY-2004; 2004US-0574032P.
XX
PA (MERI) MERCK & CO INC.
XX
PI Anderson AS;
XX
DR WPI; 2006-020409/02.
XX
PT polypeptide immunogen useful for inducing protective immunity against
PT Staphylococcus aureus comprises amino acid sequence, which is similar to
PT derivative of Staphylococcus aureus polypeptide.
XX
PS Example 1; SEQ ID NO 3; 28pp; English.
XX
CC This sequence represents a polypeptide related to the immunogen, ORF0826.
CC ORF0826 shares a high degree of homology with S. epidermidis secreted
CC antigen Saa. This immunogen can induce protective immunity against S.
CC aureus. The immunogen is administered at a dosage of 1 microg-1 mg,
CC subcutaneously, intramuscularly or mucosally. It is used for inducing
CC protective immune response in a patient e.g. human for treating
CC prophylactically against Staphylococcus aureus infection. The polypeptide
CC immunogen provides protective immunity against Staphylococcus aureus.
CC This sequence shares 85% homology to ORF0826 but is excluded from the
CC scope of the invention.
XX
SQ Sequence 162 AA;
XX
Query Match 2.3%; Score 13; DB 10; Length 162;
Best Local Similarity 100.0%; Pred. No. 0.00087;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 392 QTTTTTTTTTTS 404
Db |||||||
40 QTTTTTTTTTTS 52
XX
RESULT 8
ABJ18927
ID ABJ18927 standard; protein; 166 AA.
XX
AC ABJ18927;
XX
DT 06-MAR-2003 (first entry)
XX
DE Pathogen specific antigen related staphylococcal protein SEQ ID No 73.
XX
KW Antibacterial; virucide; fungicide; protozoacide; cytostatic; anti-HIV;
KW hyperimmune; serum-reactive; antigen; pathogen; tumour; allergen;
KW auto-immunity; vaccine; staphylococcal infection; antibody; cancer;
KW autoimmune disease; HIV; hepatitis.
XX
OS Staphylococcus sp.
XX
FN WO200259148-A2.
XX
PD 01-AUG-2002.
XX
PE 21-JAN-2002; 2002WO-BP000546.
XX
PR 26-JAN-2001; 2001AT-00000130.
XX
PA (CIST-) CISTEM BIOTECHNOLOGIES GMBH.
XX
PI Meinke A, Nagy E, Von Ahnen U, Klade C, Henics T, Zauner W;
PI Minh DB, Vytvytska O, Etz H, Dryla A, Weichhart T, Hafner M,
PI Tempelmeier B;
XX
DR WPI; 2003-075410/07.
XX

PT Identifying, isolating and producing hyperimmune serum-reactive antigens
PT from a pathogen, for preparing vaccine or medicament for treating or
PT preventing e.g. staphylococcal infections, comprises providing antibody
PT preparation.
XX
PS Claim 21; Page 157; 252pp; English.
XX
CC The invention relates to a novel method for identifying, isolating and
CC producing hyperimmune serum-reactive antigens from a pathogen, tumour,
CC allergen, a tissue or host prone to auto-immunity, where the antigens are
CC used in a vaccine, comprises providing antibody preparation from a plasma
CC pool of a type of animal, or individual sera with antibodies against the
CC specific pathogen, tumour, allergen, tissue or host prone to auto-
CC immunity. The hyperimmune serum-reactive antigens comprising any of the
CC 62 sequences of 53-2261 amino acids fully defined in the specification,
CC or their hyperimmune fragments are useful for the manufacture of a
CC pharmaceutical preparation, particularly a vaccine against staphylococcal
CC infections or colonisation against S. aureus or S. epidermidis. The
CC preparation of antibodies is useful for the manufacture of a medicament
CC for treating or preventing staphylococcal infections or colonisation
CC against S. aureus or S. epidermidis. The antibody preparations may also
CC be used for diagnostic and imaging purposes. Other conditions that can be
CC treated include cancer, autoimmune diseases or infections caused by viral
CC (e.g. HIV, hepatitis A, B or C), fungal or protozoan pathogens. This
CC sequence represents a staphylococcal protein relating to the method for
CC identifying and producing pathogen specific antigens of the invention
XX
SQ Sequence 166 AA;
XX
Query Match 2.3%; Score 13; DB 6; Length 166;
Best Local Similarity 100.0%; Pred. No. 0.00089;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 392 QTTTTTTTTTTS 404
Db |||||||
43 QTTTTTTTTTTS 55
XX
RESULT 9
ABM71151
ID ABM71151 standard; protein; 166 AA.
XX
AC ABM71151;
XX
DT 20-NOV-2003 (first entry)
XX
DE Staphylococcus aureus protein #391.
XX
KW Antibacterial; vaccine; gene therapy; infection; sepsis; diagnosis;
KW enzymatic assay; antibiotic target.
XX
OS Staphylococcus aureus.
XX
PN WO200294868-A2.
XX
PD 28-NOV-2002.
XX
PF 27-MAR-2002; 2002WO-IB002637.
XX
PR 27-MAR-2001; 2001GB-00007661.
XX
PA (CHIR-) CHIRON SPA.
XX
PI Masignani V, Mora M, Scarselli M;
XX
DR WPI; 2003-120786/11.
DR N-PSDB; ACF72711.
XX
PT New Staphylococcus aureus protein, useful as a vaccine for treating or
PT preventing Staphylococcal infection, specifically an infection caused by
PT S. aureus, e.g. sepsis.
XX
PS Claim 1; SEQ ID NO 782; 49pp; English.
XX

XX The invention relates to novel genes and encoded proteins from
 CC Staphylococcus aureus. A composition comprising the S. aureus protein, a
 CC nucleic acid encoding the protein, or an antibody to the protein, is
 CC useful as a pharmaceutical, particularly as a vaccine for treating or
 CC preventing infection due to Staphylococcus bacteria, specifically an
 CC infection caused by S. aureus. The composition is particularly useful for
 CC treating or preventing sepsis in a patient. The composition can also be
 CC used for diagnostics. The protein is also used in an assay for enzymatic
 CC studies and as a target for antibiotics. This sequence represents one of
 CC the novel S. aureus proteins of the invention
 CC
 SQ Sequence 166 AA;

Query Match 2.3%; Score 13; DB 6; Length 166;
 Best Local Similarity 100.0%; Pred. No. 0.00089;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 392 QTTTTTTTTTTS 404
 |||||
 DB 43 QTTTTTTTTTTS 55

RESULT 10
 AEE26470
 XX AEE26470 standard; protein; 166 AA.

AC AEE26470;
 XX
 XX
 DT 09-FEB-2006 (first entry)

DE ORF0826 immunogen related sequence #2.

XX Vaccine; Gene-therapy; Antibacterial; immunogen; ORF0826; Ssa;
 KM Staphylococcus aureus; infection.

XX Staphylococcus aureus; methicillin resistant strain.
 OS Synthetic.

XX WO2005115113-A2.

XX 08-DEC-2005.

XX 20-MAY-2005; 2005WO-US017835.

XX 25-MAY-2004; 2004US-0574032P.

XX (MERI) MERCK & CO INC.

XX Anderson AS;

XX WPI; 2006-020409/02.

XX Polypeptide immunogen useful for inducing protective immunity against
 PT Staphylococcus aureus comprises amino acid sequence, which is similar to
 PT derivative of Staphylococcus aureus polypeptide.

XX Example 1; SEQ ID NO 4; 28pp; English.

XX This sequence represents a polypeptide related to the immunogen, ORF0826.
 CC ORF0826 shares a high degree of homology with S. epidermis secreted
 CC antigen Ssa. This immunogen can induce protective immunity against S.
 CC aureus. The immunogen is administered at a dosage of 1 microg-1 mg,
 CC subcutaneously, intramuscularly or mucosally. It is used for inducing
 CC protective immune response in a patient e.g. human for treating
 CC prophylactically against Staphylococcus aureus infection. The polypeptide
 CC immunogen provides protective immunity against Staphylococcus aureus.
 CC This sequence shares 85% homology to ORF0826 but is excluded from the
 CC scope of the invention.

XX Sequence 166 AA;

Query Match 2.3%; Score 13; DB 10; Length 166;

Best Local Similarity 100.0%; Pred. No. 0.00089;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 392 QTTTTTTTTTTS 404
 |||||
 DB 43 QTTTTTTTTTTS 55

RESULT 11
 AEE26471
 XX AEE26471 standard; protein; 166 AA.

XX AEE26471;

XX 09-FEB-2006 (first entry)

DE ORF0826 immunogen related sequence #3.

XX Vaccine; Gene-therapy; Antibacterial; immunogen; ORF0826; Ssa;
 KM Staphylococcus aureus; infection.

XX Staphylococcus aureus; methicillin resistant strain.
 OS Synthetic.

XX WO2005115113-A2.

XX 08-DEC-2005.

XX 20-MAY-2005; 2005WO-US017835.

XX 25-MAY-2004; 2004US-0574032P.

XX (MERI) MERCK & CO INC.

XX Anderson AS;

XX WPI; 2006-020409/02.

XX Polypeptide immunogen useful for inducing protective immunity against
 PT Staphylococcus aureus comprises amino acid sequence, which is similar to
 PT derivative of Staphylococcus aureus polypeptide.

XX Example 1; SEQ ID NO 5; 28pp; English.

XX This sequence represents a polypeptide related to the immunogen, ORF0826.
 CC ORF0826 shares a high degree of homology with S. epidermis secreted
 CC antigen Ssa. This immunogen can induce protective immunity against S.
 CC aureus. The immunogen is administered at a dosage of 1 microg-1 mg,
 CC subcutaneously, intramuscularly or mucosally. It is used for inducing
 CC protective immune response in a patient e.g. human for treating
 CC prophylactically against Staphylococcus aureus infection. The polypeptide
 CC immunogen provides protective immunity against Staphylococcus aureus.
 CC This sequence shares 85% homology to ORF0826 but is excluded from the
 CC scope of the invention.

XX Sequence 166 AA;

Query Match 2.3%; Score 13; DB 10; Length 166;
 Best Local Similarity 100.0%; Pred. No. 0.00089;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 392 QTTTTTTTTTTS 404
 |||||
 DB 43 QTTTTTTTTTTS 55

RESULT 12
 AEE26472
 XX AEE26472 standard; protein; 166 AA.

XX AEE26472;

XX 09-FEB-2006 (first entry)

```

XX ORF0826 immunogen related sequence #4.
DE Vaccine; Gene-therapy; Antibacterial; immunogen; ORF0826; Ssa;
XX Staphylococcus aureus; infection.
KM Staphylococcus aureus; infection.
XX Staphylococcus aureus; methicillin resistant strain.
OS Synthetic.
XX WO2005115113-A2.
PN 08-DEC-2005.
XX 20-MAY-2005; 2005WO-US017835.
XX 25-MAY-2004; 2004US-0574032P.
PR (MERI ) MERCK & CO INC.
XX Anderson AS;
XX WPI; 2006-020409/02.
DR Polypeptide immunogen useful for inducing protective immunity against
XX Staphylococcus aureus comprises amino acid sequence, which is similar to
XX derivative of Staphylococcus aureus polypeptide.
PS Example 1; SEQ ID NO 6; 28bp; English.
XX This sequence represents a polypeptide related to the immunogen, ORF0826.
XX ORF0826 shares a high degree of homology with S. epidermis secreted
XX antigen Ssa. This immunogen can induce protective immunity against S.
XX aureus. The immunogen is administered at a dosage of 1 microg-1 mg,
XX subcutaneously, intramuscularly or mucosally. It is used for inducing
XX protective immune response in a patient e.g. human for treating
XX prophylactically against Staphylococcus aureus infection. The polypeptide
XX immunogen provides protective immunity against Staphylococcus aureus.
XX This sequence was included for comparison.
SQ Sequence 166 AA;

Query Match      2.3%; Score 13; DB 10; Length 166;
Best Local Similarity 100.0%; Pred. No. 0.00089;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 392 OTTTTTTTTTTTS 404
   |||||||
DB 43 OTTTTTTTTTTTS 55

RESULT 13
AEE26467
ID AEE26467 standard; protein; 167 AA.
XX
XX AEE26467;
AC 09-FEB-2006 (first entry)
XX
XX ORF0826 immunogen.
DE Vaccine; Gene-therapy; Antibacterial; immunogen; ORF0826; Ssa;
XX Staphylococcus aureus; infection.
OS Staphylococcus aureus.
XX WO2005115113-A2.
XX 08-DEC-2005.
XX 20-MAY-2005; 2005WO-US017835.
XX 25-MAY-2004; 2004US-0574032P.
XX
XX

```

```

PA (MERI ) MERCK & CO INC.
XX Anderson AS;
PI WPI; 2006-020409/02.
XX N-PSDB; AEE26475.
DR Polypeptide immunogen useful for inducing protective immunity against
XX Staphylococcus aureus comprises amino acid sequence, which is similar to
XX derivative of Staphylococcus aureus polypeptide.
PS Claim 1; SEQ ID NO 1; 28bp; English.
XX This sequence represents a polypeptide immunogen, ORF0826. ORF0826 shares
XX a high degree of homology with S. epidermis secreted antigen Ssa. This
XX immunogen can induce protective immunity against S. aureus. The immunogen
XX is administered at a dosage of 1 microg-1 mg, subcutaneously,
XX intramuscularly or mucosally. It is used for inducing protective immune
XX response in a patient e.g. human for treating prophylactically against
XX Staphylococcus aureus infection. The polypeptide immunogen provides
XX protective immunity against Staphylococcus aureus.
SQ Sequence 167 AA;

Query Match      2.3%; Score 13; DB 10; Length 167;
Best Local Similarity 100.0%; Pred. No. 0.0009;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 392 OTTTTTTTTTTTS 404
   |||||||
DB 44 OTTTTTTTTTTTS 56

RESULT 14
AEE26468
ID AEE26468 standard; protein; 211 AA.
XX
XX AEE26468;
AC 09-FEB-2006 (first entry)
XX
XX Full length ORF0826 immunogen.
DE Vaccine; Gene-therapy; Antibacterial; immunogen; ORF0826; Ssa;
XX Staphylococcus aureus; infection.
KM Staphylococcus aureus.
OS Staphylococcus aureus.
XX WO2005115113-A2.
XX 08-DEC-2005.
XX 20-MAY-2005; 2005WO-US017835.
XX 25-MAY-2004; 2004US-0574032P.
XX (MERI ) MERCK & CO INC.
XX Anderson AS;
XX WPI; 2006-020409/02.
XX N-PSDB; AEE26474.
DR Polypeptide immunogen useful for inducing protective immunity against
XX Staphylococcus aureus comprises amino acid sequence, which is similar to
XX derivative of Staphylococcus aureus polypeptide.
PS Example 1; SEQ ID NO 2; 28bp; English.
XX This sequence represents a polypeptide m from which the immunogen,
XX ORF0826, is derived. ORF0826 shares a high degree of homology with S.
XX epidermis secreted antigen Ssa. This immunogen can induce protective
XX immunity against S. aureus. The immunogen is administered at a dosage of

```

CC 1 microg-1 mg, subcutaneously, intramuscularly or mucosally. It is used
 CC for inducing protective immune response in a patient e.g. human for
 CC treating prophylactically against *Staphylococcus aureus* infection. The
 CC polypeptide immunogen provides protective immunity against *Staphylococcus*
 CC *aureus*.
 CC
 XX

Sequence 211 AA;

Query Match 2.3%; Score 13; DB 10; Length 211;
 Best Local Similarity 100.0%; Pred. No. 0.0011;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 392 QTTTTTTTTTTS 404
 |||||
 DB 88 QTTTTTTTTTTS 100

RESULT 15

AEF20212
 ID AEF20212 standard; protein; 81 AA.
 XX

AC AEF20212;
 XX

DT 23-MAR-2006 (first entry)
 XX

DE *Pyrococcus furiosus* chitinase connection region SEQ ID NO:6.
 XX

KW chitinase; chitin; hydrolysis.
 XX

OS *Pyrococcus furiosus*.
 XX

PN JP2006025701-A.
 XX

PD 02-FEB-2006.
 XX

PF 16-JUL-2004; 2004JP-00209383.
 XX

PR 16-JUL-2004; 2004JP-00209383.
 XX

PA (DOKU-) DOKURITSU GYOSEI HOJIN SANGYO GIJUTSU SO.
 XX

PI Ishikawa K, Oku T;
 XX

DR MPI: 2006-121630/13.
 XX

DR N-PSDB; AEF20218.
 XX

PT Novel chitinase comprising a chitin binding domain connected to a
 PT catalytically active domain, through a direct or connecting region,
 PT useful for hydrolyzing chitin.
 XX

PS Claim 4, SEQ ID NO 6; 15pp; Japanese.
 XX

XX The invention relates to a chitinase (I) comprising a chitin binding
 CC domain (D1) connected to a catalytically active domain (D2) comprising
 CC the 289 amino acid sequence of AEF20209, or AEF20209 in which one or more
 CC amino acids being deleted, substituted or added, through a direct or a
 CC connecting region. Also described: (1) a DNA (II) encoding (1); (2) a
 CC recombinant vector (III) containing (II); (3) a transformed host (IV)
 CC containing (III); and (4) producing (1). (I) is useful for hydrolyzing
 CC chitin. (II) exhibits high activity and stability, even at high
 CC temperature. The present sequence represents a *Pyrococcus furiosus*
 CC chitinase connection region amino acid sequence, which is given in the
 CC exemplification of the present invention.
 CC
 XX

Sequence 81 AA;

Query Match 2.2%; Score 12; DB 10; Length 81;
 Best Local Similarity 100.0%; Pred. No. 0.0044;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 393 TTTTTTTTTTTS 404
 |||||
 DB 57 TTTTTTTTTTTS 68

Search completed: August 1, 2006, 22:33:43
 Job time : 241 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 1, 2006, 22:38:46 ; Search time 49 Seconds

(Without alignments)
996.777 Million cell updates/sec

Title: US-09-189-415D-11

Perfect score: 558

Sequence: 1 MPICN/GHNVNNSIPAP.....SNGAVNTSNNPAPGSHRFV 558

Scoring table: OLIGO

Searched: 650591 seqs, 87530628 residues

Word size : 8

Total number of hits satisfying chosen parameters: 269

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database :

1: Issued Patente AA:*

2: /EMC_Celerra_SIDS3/prodata/2/1aa/5 COMB.pep:*

3: /EMC_Celerra_SIDS3/prodata/2/1aa/6 COMB.pep:*

4: /EMC_Celerra_SIDS3/prodata/2/1aa/H COMB.pep:*

5: /EMC_Celerra_SIDS3/prodata/2/1aa/PCITUS COMB.pep:*

6: /EMC_Celerra_SIDS3/prodata/2/1aa/RE COMB.pep:*

7: /EMC_Celerra_SIDS3/prodata/2/1aa/backfill1st.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	12	2.2	160	2	US-09-248-796A-17879
2	12	2.2	292	2	US-09-248-796A-25055
3	12	2.2	541	2	US-09-248-796A-18318
4	12	2.2	667	2	US-09-248-796A-22880
5	12	2.2	3712	2	US-10-037-417-48
6	12	2.2	3712	2	US-10-037-417-51
7	12	2.2	3913	2	US-09-949-016-10933
8	12	2.2	4377	2	US-09-949-016-6978
9	11	2.0	13	2	US-10-029-212-10
10	11	2.0	41	2	US-09-060-767B-5
11	11	2.0	57	2	US-08-900-230-59
12	11	2.0	63	2	US-09-248-796A-23083
13	11	2.0	75	2	US-09-248-796A-25289
14	11	2.0	91	2	US-08-700-651-14
15	11	2.0	91	2	US-08-928-361B-19
16	11	2.0	91	2	US-09-588-995A-19
17	11	2.0	105	2	US-09-248-796A-22875
18	11	2.0	106	2	US-09-270-767-36192
19	11	2.0	106	2	US-09-270-767-51409
20	11	2.0	124	2	US-08-700-651-11
21	11	2.0	124	2	US-08-928-361B-16
22	11	2.0	124	2	US-09-588-995A-16
23	11	2.0	128	2	US-08-700-651-7
24	11	2.0	128	2	US-08-928-361B-12
25	11	2.0	128	2	US-09-588-995A-12
26	11	2.0	130	2	US-08-700-651-8

27	11	2.0	130	2	US-08-700-651-9	Sequence 9, Appli
28	11	2.0	130	2	US-08-928-361B-13	Sequence 13, Appli
29	11	2.0	130	2	US-08-928-361B-14	Sequence 14, Appli
30	11	2.0	130	2	US-09-588-995A-13	Sequence 13, Appli
31	11	2.0	130	2	US-09-588-995A-14	Sequence 14, Appli
32	11	2.0	137	2	US-09-248-796A-26561	Sequence 26561, A
33	11	2.0	138	2	US-08-700-651-10	Sequence 10, Appli
34	11	2.0	138	2	US-08-928-361B-15	Sequence 15, Appli
35	11	2.0	138	2	US-09-588-995A-15	Sequence 15, Appli
36	11	2.0	140	2	US-09-248-796A-23990	Sequence 23990, A
37	11	2.0	140	2	US-08-928-361B-18	Sequence 18, Appli
38	11	2.0	150	2	US-09-588-995A-18	Sequence 18, Appli
39	11	2.0	159	2	US-09-248-796A-21631	Sequence 21631, A
40	11	2.0	152	2	US-08-700-651-13	Sequence 13, Appli
41	11	2.0	175	2	US-08-700-651-12	Sequence 12, Appli
42	11	2.0	175	2	US-08-928-361B-17	Sequence 17, Appli
43	11	2.0	181	2	US-09-588-995A-17	Sequence 17, Appli
44	11	2.0	181	2	US-09-248-796A-21013	Sequence 21013, A
45	11	2.0	189	2	US-09-248-796A-22901	Sequence 22901, A
46	11	2.0	197	2	US-09-248-796A-21069	Sequence 21069, A
47	11	2.0	207	2	US-09-248-796A-16058	Sequence 16058, A
48	11	2.0	216	2	US-08-928-361B-8	Sequence 8, Appli
49	11	2.0	216	2	US-08-928-361B-27	Sequence 27, Appli
50	11	2.0	216	2	US-09-588-995A-8	Sequence 8, Appli
51	11	2.0	216	2	US-09-248-796A-17391	Sequence 17391, A
52	11	2.0	247	2	US-09-248-796A-24111	Sequence 24111, A
53	11	2.0	249	2	US-08-700-651-15	Sequence 15, Appli
54	11	2.0	249	2	US-08-928-361B-20	Sequence 20, Appli
55	11	2.0	249	2	US-09-588-995A-20	Sequence 20, Appli
56	11	2.0	288	2	US-09-216-393B-341	Sequence 341, App
57	11	2.0	288	2	US-09-248-796A-18458	Sequence 18458, A
58	11	2.0	322	2	US-09-248-796A-25420	Sequence 25420, A
59	11	2.0	331	2	US-09-270-767-43550	Sequence 43550, A
60	11	2.0	371	2	US-09-248-796A-20867	Sequence 20867, A
61	11	2.0	423	2	US-09-778-510-22	Sequence 22, Appli
62	11	2.0	439	2	US-09-248-796A-24059	Sequence 24059, A
63	11	2.0	440	2	US-09-866-028-61	Sequence 61, Appli
64	11	2.0	440	2	US-09-944-457-61	Sequence 61, Appli
65	11	2.0	440	2	US-09-945-584-61	Sequence 61, Appli
66	11	2.0	440	2	US-09-944-944-61	Sequence 61, Appli
67	11	2.0	440	2	US-09-945-587-61	Sequence 61, Appli
68	11	2.0	440	2	US-09-945-587-61	Sequence 61, Appli
69	11	2.0	440	3	US-09-944-884-61	Sequence 61, Appli
70	11	2.0	442	2	US-09-778-510-20	Sequence 20, Appli
71	11	2.0	442	2	US-09-930-803-1	Sequence 1, Appli
72	11	2.0	459	2	US-09-248-796A-18432	Sequence 18432, A
73	11	2.0	543	2	US-09-248-796A-22504	Sequence 22504, A
74	11	2.0	559	1	US-08-368-071-12	Sequence 12, Appli
75	11	2.0	559	1	US-08-458-181-12	Sequence 12, Appli
76	11	2.0	559	2	US-09-900-708-2	Sequence 2, Appli
77	11	2.0	559	5	PCT-US93-02172-12	Sequence 12, Appli
78	11	2.0	612	2	US-09-248-796A-24557	Sequence 24557, A
79	11	2.0	692	2	US-09-248-796A-18612	Sequence 18612, A
80	11	2.0	694	2	US-10-029-180-22	Sequence 22, Appli
81	11	2.0	734	2	US-09-248-796A-18305	Sequence 18305, A
82	11	2.0	887	1	US-07-867-106-3	Sequence 3, Appli
83	11	2.0	1002	2	US-09-762-724-4	Sequence 4, Appli
84	11	2.0	1060	2	US-09-248-796A-16624	Sequence 16624, A
85	11	2.0	1128	2	US-09-627-650B-11	Sequence 11, Appli
86	11	2.0	1128	2	US-09-436-063C-11	Sequence 11, Appli
87	11	2.0	1121	2	US-08-700-651-5	Sequence 5, Appli
88	11	2.0	121	2	US-08-928-361B-6	Sequence 6, Appli
89	11	2.0	121	2	US-09-588-995A-6	Sequence 6, Appli
90	11	2.0	1337	2	US-08-928-361B-5	Sequence 5, Appli
91	11	2.0	1337	2	US-09-588-995A-5	Sequence 5, Appli
92	10	1.8	25	2	US-09-060-767B-7	Sequence 7, Appli
93	10	1.8	41	2	US-09-060-767B-8	Sequence 8, Appli
94	10	1.8	57	2	US-09-060-767B-9	Sequence 9, Appli
95	10	1.8	57	2	US-09-060-767B-6	Sequence 6, Appli
96	10	1.8	65	2	US-09-248-796A-23985	Sequence 23985, A
97	10	1.8	82	2	US-09-248-796A-20742	Sequence 20742, A
98	10	1.8	84	2	US-09-248-796A-23953	Sequence 23953, A
99	10	1.8	158	2	US-09-248-796A-21630	Sequence 21630, A

100	1.8	181	2	US-09-248-796A-15978	Sequence 15978, A	173	1.6	2544	2	US-09-627-650B-3	Sequence 3, Appl1
101	1.8	210	2	US-09-252-991A-10886	Sequence 30886, A	174	1.6	2544	2	US-09-436-063C-3	Sequence 9, Appl1
102	1.8	211	2	US-09-060-767B-3	Sequence 3, Appl1	175	1.6	2601	2	US-09-627-650B-9	Sequence 9, Appl1
103	1.8	294	2	US-09-248-796A-15817	Sequence 15817, A	176	1.6	2601	2	US-09-436-063C-9	Sequence 9, Appl1
104	1.8	309	2	US-09-248-796A-20897	Sequence 20897, A	177	1.4	61	2	US-09-248-796A-22013	Sequence 22013, A
105	1.8	362	2	US-09-248-796A-14994	Sequence 14994, A	178	1.4	74	2	US-09-248-796A-27614	Sequence 27614, A
106	1.8	392	2	US-09-248-796A-16779	Sequence 16779, A	179	1.4	85	2	US-09-248-796A-21225	Sequence 21225, A
107	1.8	394	2	US-09-248-796A-23220	Sequence 23220, A	180	1.4	90	2	US-09-248-796A-23974	Sequence 23974, A
108	1.8	405	2	US-09-248-796A-23277	Sequence 23277, A	181	1.4	94	2	US-09-270-767-45952	Sequence 45952, A
109	1.8	447	2	US-09-248-796A-14299	Sequence 14299, A	182	1.4	101	2	US-09-270-767-44082	Sequence 44082, A
110	1.8	505	2	US-09-248-796A-16524	Sequence 16524, A	183	1.4	102	2	US-09-248-796A-22183	Sequence 22183, A
111	1.8	613	2	US-09-248-796A-16524	Sequence 16524, A	184	1.4	103	2	US-09-248-796A-17128	Sequence 17128, A
112	1.8	624	2	US-09-270-767-41894	Sequence 41894, A	185	1.4	105	2	US-09-248-796A-21408	Sequence 21408, A
113	1.8	663	2	US-09-248-796A-25441	Sequence 25441, A	186	1.4	123	2	US-09-248-796A-18357	Sequence 18357, A
114	1.8	666	2	US-09-248-796A-15507	Sequence 15507, A	187	1.4	123	2	US-09-248-796A-27877	Sequence 27877, A
115	1.8	788	1	US-08-194-338-12	Sequence 12, Appl	188	1.4	130	2	US-09-270-767-41135	Sequence 41135, A
116	1.8	1009	2	US-09-762-724-10	Sequence 10, Appl	189	1.4	130	2	US-09-270-767-56351	Sequence 56351, A
117	1.8	1017	2	US-09-762-724-12	Sequence 12, Appl	190	1.4	130	2	US-09-248-796A-66717	Sequence 66717, A
118	1.8	1023	2	US-09-762-724-14	Sequence 14, Appl	191	1.4	144	2	US-09-270-767-59701	Sequence 59701, A
119	1.8	1027	2	US-09-762-724-8	Sequence 8, Appl1	192	1.4	146	2	US-09-248-796A-16057	Sequence 16057, A
120	1.8	1029	2	US-09-762-724-6	Sequence 6, Appl1	193	1.4	162	2	US-09-270-767-43570	Sequence 43570, A
121	1.6	13	2	US-10-029-212-11	Sequence 11, Appl	194	1.4	164	2	US-09-543-681A-8262	Sequence 8262, Ap
122	1.6	63	2	US-09-248-796A-21124	Sequence 21124, A	195	1.4	166	2	US-09-248-796A-14829	Sequence 14829, A
123	1.6	68	2	US-09-248-796A-21701	Sequence 21701, A	196	1.4	174	2	US-09-270-767-44506	Sequence 44506, A
124	1.6	102	2	US-09-248-796A-27894	Sequence 27894, A	197	1.4	175	2	US-09-640-211A-2223	Sequence 2223, Ap
125	1.6	102	2	US-09-513-999C-6366	Sequence 6366, Ap	198	1.4	192	2	US-09-711-334-2	Sequence 21055, A
126	1.6	133	2	US-09-370-767-39290	Sequence 39290, A	199	1.4	199	2	US-09-248-796A-21055	Sequence 21314, A
127	1.6	133	2	US-09-370-767-34507	Sequence 54507, A	200	1.4	219	2	US-09-248-796A-15839	Sequence 15839, A
128	1.6	153	2	US-09-248-796A-14255	Sequence 14255, A	201	1.4	248	2	US-09-248-796A-21208	Sequence 21208, A
129	1.6	159	2	US-09-248-796A-16802	Sequence 16802, A	202	1.4	258	2	US-09-248-796A-17309	Sequence 17309, A
130	1.6	162	2	US-09-248-796A-16630	Sequence 16630, A	203	1.4	266	2	US-09-270-767-33574	Sequence 32574, A
131	1.6	168	2	US-09-746-801A-63	Sequence 63, Appl	204	1.4	268	2	US-09-270-767-47791	Sequence 47791, A
132	1.6	168	2	US-10-719-885-63	Sequence 63, Appl	205	1.4	268	2	US-09-248-796A-20340	Sequence 20340, A
133	1.6	175	2	US-09-270-767-33553	Sequence 33553, A	206	1.4	269	2	US-09-248-796A-22276	Sequence 22276, A
134	1.6	175	2	US-09-270-767-48770	Sequence 48770, A	207	1.4	277	2	US-08-842-445-34	Sequence 34, Appl
135	1.6	186	2	US-09-248-796A-15844	Sequence 15844, A	208	1.4	277	2	US-09-186-188B-34	Sequence 34, Appl
136	1.6	195	2	US-09-248-796A-21930	Sequence 21930, A	209	1.4	277	2	US-09-842-445-34	Sequence 34, Appl
137	1.6	218	2	US-09-248-796A-15306	Sequence 15306, A	210	1.4	283	2	US-09-186-188B-34	Sequence 34, Appl
138	1.6	224	2	US-09-248-796A-18758	Sequence 18758, A	211	1.4	285	2	US-09-248-796A-15972	Sequence 15972, A
139	1.6	229	2	US-09-248-796A-11127	Sequence 17127, A	212	1.4	290	2	US-09-248-796A-21613	Sequence 21613, A
140	1.6	240	2	US-09-248-796A-24008	Sequence 24008, A	213	1.4	313	2	US-09-270-767-31793	Sequence 37193, A
141	1.6	262	2	US-09-248-796A-21832	Sequence 21832, A	214	1.4	319	2	US-09-270-767-5410	Sequence 52410, A
142	1.6	263	2	US-09-248-796A-14400	Sequence 14400, A	215	1.4	336	2	US-09-270-767-43343	Sequence 43343, A
143	1.6	277	2	US-09-248-796A-17557	Sequence 17557, A	216	1.4	339	2	US-09-686-583B-40	Sequence 40, Appl
144	1.6	299	2	US-09-248-796A-23243	Sequence 23243, A	217	1.4	357	1	US-08-078-683A-8	Sequence 8, Appl1
145	1.6	308	2	US-09-248-796A-21293	Sequence 21293, A	218	1.4	357	1	US-08-471-970A-8	Sequence 8, Appl1
146	1.6	333	2	US-09-248-796A-25847	Sequence 25847, A	219	1.4	357	2	US-09-723-677B-8	Sequence 20, Appl
147	1.6	347	2	US-09-248-796A-16323	Sequence 16323, A	220	1.4	366	2	US-09-377-285B-20	Sequence 20, Appl
148	1.6	361	2	US-09-248-796A-19234	Sequence 19234, A	221	1.4	366	2	US-10-192-381-20	Sequence 4, Appl1
149	1.6	373	2	US-09-248-796A-18953	Sequence 18953, A	222	1.4	366	2	US-09-404-296B-4	Sequence 48, Appl
150	1.6	383	2	US-09-248-796A-23916	Sequence 23916, A	223	1.4	373	2	US-09-270-767-44277	Sequence 23864, A
151	1.6	436	2	US-09-270-767-41918	Sequence 41918, A	224	1.4	381	2	US-09-686-583B-48	Sequence 22604, A
152	1.6	456	2	US-09-270-767-46336	Sequence 46336, A	225	1.4	384	2	US-09-248-796A-23584	Sequence 23584, A
153	1.6	476	2	US-09-248-796A-18994	Sequence 18994, A	226	1.4	393	2	US-09-248-796A-22604	Sequence 43, Appl
154	1.6	492	2	US-09-657-013-66	Sequence 66, Appl	227	1.4	397	2	US-09-686-583B-43	Sequence 21599, A
155	1.6	518	2	US-09-248-796A-19845	Sequence 19845, A	228	1.4	424	2	US-09-265-585C-149	Sequence 149, App
156	1.6	553	1	US-08-651-572-2	Sequence 2, Appl1	229	1.4	426	2	US-09-248-796A-25086	Sequence 25086, A
157	1.6	553	2	US-09-066-544-2	Sequence 2, Appl1	230	1.4	439	1	US-08-507-431-4	Sequence 4, Appl1
158	1.6	553	2	US-08-951-086-2	Sequence 2, Appl1	231	1.4	439	1	US-08-700-546-2	Sequence 2, Appl1
159	1.6	576	2	US-09-430-669-2	Sequence 20509, A	232	1.4	439	1	US-08-902-655A-4	Sequence 4, Appl1
160	1.6	593	2	US-09-248-796A-20509	Sequence 19340, A	233	1.4	439	1	US-09-116-622-4	Sequence 4, Appl1
161	1.6	1404	1	US-08-400-159-2	Sequence 2, Appl1	234	1.4	439	2	US-09-219-277-4	Sequence 4, Appl1
162	1.6	1404	2	US-08-611-729A-2	Sequence 2, Appl1	235	1.4	439	2	US-09-599-661-4	Sequence 1, Appl1
163	1.6	1404	2	US-09-195-524-2	Sequence 1, Appl1	236	1.4	476	2	US-09-382-086-1	Sequence 15175, A
164	1.6	1652	2	US-09-627-650B-1	Sequence 1, Appl1	237	1.4	476	2	US-09-248-796A-15175	Sequence 22382, A
165	1.6	1652	2	US-09-436-063C-1	Sequence 1, Appl1	238	1.4	497	2	US-09-248-796A-18965	Sequence 18965, A
166	1.6	1798	2	US-09-370-767-60233	Sequence 60233, A	239	1.4	521	2	US-09-248-796A-23363	Sequence 23363, A
167	1.6	1917	2	US-09-627-650B-5	Sequence 5, Appl1	240	1.4	525	2	US-10-101-464A-63	Sequence 63, App
168	1.6	1917	2	US-09-436-063C-5	Sequence 44775, A	241	1.4	588	2	US-09-265-585C-34	Sequence 34, Appl
169	1.6	2410	2	US-09-270-767-44775	Sequence 7, Appl1	242	1.4				
170	1.6	2508	2	US-09-627-650B-7	Sequence 7, Appl1	243	1.4				
171	1.6	2508	2	US-09-436-063C-7	Sequence 7, Appl1	244	1.4				
172	1.6	2508	2	US-09-436-063C-7	Sequence 7, Appl1	245	1.4				

```
246 8 1.4 618 2 US-09-248-796A-19273 Sequence 19273, A
247 8 1.4 620 2 US-10-087-167-137 Sequence 137, App
248 8 1.4 623 2 US-09-949-016-6530 Sequence 6530, App
249 8 1.4 640 2 US-09-252-991A-20343 Sequence 20343, A
250 8 1.4 668 2 US-09-248-796A-19350 Sequence 19350, A
251 8 1.4 672 2 US-09-336-115C-10 Sequence 10, App1
252 8 1.4 684 2 US-09-303-518D-721 Sequence 721, App
253 8 1.4 692 2 US-09-248-796A-17390 Sequence 17390, A
254 8 1.4 693 2 US-09-949-016-7806 Sequence 7806, App
255 8 1.4 721 2 US-09-185-160-11 Sequence 11, App1
256 8 1.4 747 2 US-09-824-574-6 Sequence 6, App1
257 8 1.4 752 2 US-10-104-047-1975 Sequence 1975, App
258 8 1.4 780 2 US-10-148-806-35 Sequence 35, App1
259* 8 1.4 793 2 US-10-055-364-55 Sequence 55, App1
260 8 1.4 801 1 US-07-906-349A-6 Sequence 6, App1
261 8 1.4 830 2 US-10-055-364-45 Sequence 45, App1
262 8 1.4 849 2 US-09-949-016-9522 Sequence 9522, App
263 8 1.4 887 2 US-09-077-940A-2 Sequence 2, App1
264 8 1.4 892 2 US-09-248-796A-19224 Sequence 19224, A
265 8 1.4 1495 2 US-09-697-898-3 Sequence 3, App1
266 8 1.4 1512 2 US-09-697-898-2 Sequence 2, App1
267 8 1.4 1648 2 US-09-248-796A-14286 Sequence 14286, A
268 8 1.4 1817 2 US-09-004-838-125 Sequence 125, App
269 8 1.4 5179 2 US-09-538-092-1258 Sequence 1258, App
```

ALIGNMENTS

```
RESULT 1
US-09-248-796A-17879
; Sequence 17879, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 17879
; LENGTH: 160
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-17879

Query Match      2.2%; Score 12; DB 2; Length 160;
Best Local Similarity 100.0%; Pred. No. 0.0033;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      392 OTTTTTTTTTT 403
DB      2 OTTTTTTTTTT 13

RESULT 2
US-09-248-796A-25055
; Sequence 25055, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-02-13
```

```
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 25055
; LENGTH: 292
; TYPE: PRT
; ORGANISM: Candida albicans
; NAME/KEY: UNSURE
; LOCATION: (287), (288), (289)
; OTHER INFORMATION: Identity of amino acid sequences at the above locations are unknown
US-09-248-796A-25055
```

```
Query Match      2.2%; Score 12; DB 2; Length 292;
Best Local Similarity 100.0%; Pred. No. 0.0059;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      393 TTTTTTTTTT 404
DB      97 TTTTTTTTTT 108
```

```
RESULT 3
US-09-248-796A-18318
; Sequence 18318, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 18318
; LENGTH: 541
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-18318

Query Match      2.2%; Score 12; DB 2; Length 541;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      393 TTTTTTTTTT 404
DB      207 TTTTTTTTTT 218

RESULT 4
US-09-248-796A-22880
; Sequence 22880, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 22880
; LENGTH: 667
; TYPE: PRT
; ORGANISM: Candida albicans
```

```
FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (667)
; OTHER INFORMATION: Identity of amino acid sequences at the above locations are unknown
US-09-248-796A-22880
```

```
Query Match      2.2%; Score 12; DB 2; Length 667;
Best Local Similarity 100.0%; Pred. No. 0.013;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      393 TTTTTTTTTTS 404
          |||||
Db      585 TTTTTTTTTTS 596
```

```
RESULT 5
US-10-037-417-48
; Sequence 48, Application US/10037417
; Patent No. 6903201
; GENERAL INFORMATION:
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Alsobrook II, John P
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Liu, Xiaohong
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Paturajan, Meera
; APPLICANT: Grosse, William M
; APPLICANT: Lepley, Denise M
; APPLICANT: Burgess, Catherine E
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Li, Li
; APPLICANT: Gorman, Linda
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Sciore, Paul
; APPLICANT: Ellerman, Karen
; APPLICANT: Malysankar, Uriel M
; APPLICANT: Rothenberg, Mark
; APPLICANT: Stone, David J
; APPLICANT: Boldog, Ferenc L
; APPLICANT: Guo, Xiaojia
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Anderson, David W
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Taupier Jr, Raymond J
; APPLICANT: Miller, Charles E
; APPLICANT: Eisen, Andrew J
; FILE REFERENCE: 21402-235
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; CURRENT APPLICATION NUMBER: US/10/037,417
; PRIOR FILING DATE: 2002-09-20
; PRIOR APPLICATION NUMBER: 60/260,018
; PRIOR FILING DATE: 2001-01-05
; PRIOR APPLICATION NUMBER: 60/260,360
; PRIOR FILING DATE: 2001-01-08
; PRIOR APPLICATION NUMBER: 60/272,411
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/272,817
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: 60/291,186
; PRIOR FILING DATE: 2001-05-15
; PRIOR APPLICATION NUMBER: 60/303,231
; PRIOR FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 60/305,060
; PRIOR FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: 60/318,405
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/318,700
; PRIOR FILING DATE: 2001-09-12
; NUMBER OF SEQ ID NOS: 227
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 48
; LENGTH: 3712
; TYPE: PRT
```

```
; ORGANISM: Drosophila melanogaster
US-10-037-417-48
```

```
Query Match      2.2%; Score 12; DB 2; Length 3712;
Best Local Similarity 100.0%; Pred. No. 0.067;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      393 TTTTTTTTTTS 404
          |||||
Db      3273 TTTTTTTTTTS 3284
```

```
RESULT 6
US-10-037-417-51
; Sequence 51, Application US/10037417
; Patent No. 6903201
; GENERAL INFORMATION:
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Alsobrook II, John P
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Liu, Xiaohong
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Paturajan, Meera
; APPLICANT: Grosse, William M
; APPLICANT: Lepley, Denise M
; APPLICANT: Burgess, Catherine E
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Li, Li
; APPLICANT: Gorman, Linda
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Sciore, Paul
; APPLICANT: Ellerman, Karen
; APPLICANT: Malysankar, Uriel M
; APPLICANT: Rothenberg, Mark
; APPLICANT: Stone, David J
; APPLICANT: Boldog, Ferenc L
; APPLICANT: Guo, Xiaojia
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Anderson, David W
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Taupier Jr, Raymond J
; APPLICANT: Miller, Charles E
; APPLICANT: Eisen, Andrew J
; FILE REFERENCE: 21402-235
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; CURRENT APPLICATION NUMBER: US/10/037,417
; PRIOR FILING DATE: 2002-09-20
; PRIOR APPLICATION NUMBER: 60/260,018
; PRIOR FILING DATE: 2001-01-05
; PRIOR APPLICATION NUMBER: 60/260,360
; PRIOR FILING DATE: 2001-01-08
; PRIOR APPLICATION NUMBER: 60/272,411
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/272,817
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: 60/291,186
; PRIOR FILING DATE: 2001-05-15
; PRIOR APPLICATION NUMBER: 60/303,231
; PRIOR FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 60/305,060
; PRIOR FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: 60/318,405
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/318,700
; PRIOR FILING DATE: 2001-09-12
; NUMBER OF SEQ ID NOS: 227
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 51
; LENGTH: 3712
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-10-037-417-51
```

Query Match 2.2%; Score 12; DB 2; Length 3712;
Best Local Similarity 100.0%; Pred. No. 0.067;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 393 TTTT TTTT TTTT 404
|||||
Db 3273 TTTT TTTT TTTT 3284

RESULT 7
US-09-949-016-10933
; Sequence 10933, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10933
; LENGTH: 3913
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-10933

Query Match 2.2%; Score 12; DB 2; Length 3913;
Best Local Similarity 100.0%; Pred. No. 0.07;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

QY 393 TTTT TTTT TTTT 404
|||||
Db 3504 TTTT TTTT TTTT 3515

RESULT 8
US-09-949-016-6978
; Sequence 6978, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6978
; LENGTH: 4377
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6978

Query Match 2.2%; Score 12; DB 2; Length 4377;
Best Local Similarity 100.0%; Pred. No. 0.078;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 393 TTTT TTTT TTTT 404

Db 3968 TTTT TTTT TTTT 3979
|||||

RESULT 9
US-10-029-212-10
; Sequence 10, Application US/10029212
; Patent No. 6770748
; GENERAL INFORMATION:
; APPLICANT: IMANISHI, Takeshi
; APPLICANT: OBIKA, Satoshi
; TITLE OF INVENTION: NOVEL BICYCLONUCLEOSIDE AND OLIGONUCLEOTIDE ANALOGUE
; FILE REFERENCE: IMANISHI 28
; CURRENT APPLICATION NUMBER: US/10/029,212
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: 09/904,567
; PRIOR FILING DATE: 2001-07-17
; PRIOR APPLICATION NUMBER: 09/380,638
; PRIOR FILING DATE: 1999-09-07
; PRIOR APPLICATION NUMBER: PCT/JP98/00945
; PRIOR FILING DATE: 1998-03-09
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 10
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: synthetic
US-10-029-212-10

Query Match 2.0%; Score 11; DB 2; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.0028;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 393 TTTT TTTT TTTT 403
|||||
Db 2 TTTT TTTT TTTT 12

RESULT 10
US-09-060-767B-5
; Sequence 5, Application US/09060767B
; Patent No. 6720152
; GENERAL INFORMATION:
; APPLICANT: Weil, Gary
; APPLICANT: Chandrasekar, Ramaswamy
; TITLE OF INVENTION: Diagnosis of Histoplasmosis Using Antigens Specific for
; FILE REFERENCE: BUCH 9986
; CURRENT APPLICATION NUMBER: US/09/060,767B
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/043,332
; PRIOR FILING DATE: 1997-04-15
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 5
; LENGTH: 41
; TYPE: PRT
; ORGANISM: Leishmania
US-09-060-767B-5

Query Match 2.0%; Score 11; DB 2; Length 41;
Best Local Similarity 100.0%; Pred. No. 0.0083;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 393 TTTT TTTT TTTT 403
|||||
Db 3 TTTT TTTT TTTT 13

RESULT 11
US-08-900-230-59

```
/ Sequence 59, Application US/08900230
/ Patent No. 6329197
/ GENERAL INFORMATION:
/ APPLICANT: Baird, Jonathan A.
/ TITLE OF INVENTION: DNA ENCODING GALANN GALR3 RECEPTORS AND
/ TITLE OF INVENTION: US9248796A
/ NUMBER OF SEQUENCES: 59
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Cooper & Dunham LLP
/ STREET: 1185 Avenue of the Americas
/ CITY: New York
/ STATE: New York
/ COUNTRY: U.S.A.
/ ZIP: 11036
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/900,230
/ FILING DATE: 23-JUL-1997
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: White, John P.
/ REGISTRATION NUMBER: 28,678
/ REFERENCE/DOCKET NUMBER: 52241-C/JPW/ADM
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 212-278-0400
/ TELEFAX: 212-391-0525
/ INFORMATION FOR SEQ ID NO: 59:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 57 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE:
/ HYPOTHEICAL: NO
/ ANTI-SENSE: NO
/ US-08-900-230-59

Query Match          2.0%; Score 11; DB 2; Length 57;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      393 TTTT TTTT TTTT 403
Db      1 TTTT TTTT TTTT 11

RESULT 12
US-09-248-796A-23083
/ Sequence 23083, Application US/09248796A
/ Patent No. 6747137
/ GENERAL INFORMATION:
/ APPLICANT: Keith Weinstein et al
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
/ TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
/ FILE REFERENCE: 107196.132
/ CURRENT APPLICATION NUMBER: US/09/248,796A
/ CURRENT FILING DATE: 1999-02-12
/ PRIOR APPLICATION NUMBER: US 60/074,725
/ PRIOR FILING DATE: 1998-02-13
/ PRIOR APPLICATION NUMBER: US 60/096,409
/ PRIOR FILING DATE: 1998-08-13
/ NUMBER OF SEQ ID NOS: 28208
/ SEQ ID NO 23083
/ LENGTH: 63
/ TYPE: PRT
/ ORGANISM: Candida albicans
/ US-09-248-796A-23083

Query Match          2.0%; Score 11; DB 2; Length 63;
```

```
Best Local Similarity 100.0%; Pred. No. 0.012;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      393 TTTT TTTT TTTT 403
Db      37 TTTT TTTT TTTT 47

RESULT 13
US-09-248-796A-25289
/ Sequence 25289, Application US/09248796A
/ Patent No. 6747137
/ GENERAL INFORMATION:
/ APPLICANT: Keith Weinstein et al
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
/ TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
/ FILE REFERENCE: 107196.132
/ CURRENT APPLICATION NUMBER: US/09/248,796A
/ CURRENT FILING DATE: 1999-02-12
/ PRIOR APPLICATION NUMBER: US 60/074,725
/ PRIOR FILING DATE: 1998-02-13
/ PRIOR APPLICATION NUMBER: US 60/096,409
/ PRIOR FILING DATE: 1998-08-13
/ NUMBER OF SEQ ID NOS: 28208
/ SEQ ID NO 25289
/ LENGTH: 75
/ TYPE: PRT
/ ORGANISM: Candida albicans
/ US-09-248-796A-25289

Query Match          2.0%; Score 11; DB 2; Length 75;
Best Local Similarity 100.0%; Pred. No. 0.015;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      393 TTTT TTTT TTTT 403
Db      7 TTTT TTTT TTTT 17

RESULT 14
US-08-700-651-14
/ Sequence 14, Application US/08700651B
/ Patent No. 6015882
/ GENERAL INFORMATION:
/ APPLICANT: PETERSEN, CAROLYN
/ APPLICANT: LEBCH, JAMES
/ APPLICANT: NELSON, RICHARD, C.
/ APPLICANT: GUT, JIRI
/ TITLE OF INVENTION: VACCINES, ANTIBODIES, PROTEINS, GLYCOPROTEINS, DNAs AND RNAs
/ TITLE OF INVENTION: FOR PROPHYLAXIS AND TREATMENT OF Cryptosporidium parvum
/ FILE REFERENCE: 480.19-4(HV)
/ CURRENT APPLICATION NUMBER: US/08/700,651B
/ CURRENT FILING DATE: 1997-08-14
/ EARLIER APPLICATION NUMBER: 08/415,751
/ EARLIER FILING DATE: 1995-04-03
/ NUMBER OF SEQ ID NOS: 15
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 14
/ LENGTH: 91
/ TYPE: PRT
/ ORGANISM: Cryptosporidium parvum
/ FEATURE:
/ OTHER INFORMATION: mutant/variant of SEQ ID NO:5
/ US-08-700-651-14

Query Match          2.0%; Score 11; DB 2; Length 91;
Best Local Similarity 100.0%; Pred. No. 0.018;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      393 TTTT TTTT TTTT 403
Db      18 TTTT TTTT TTTT 28
```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 1, 2006, 22:50:21 ; Search time 181 Seconds
(without alignments)
1428.033 Million cell updates/sec

Title: US-09-189-415d-11

Perfect score: 558

Sequence: 1 MEIGNLGNPNVNSIPPAP.....SNGAVNTSNPPAGSHRFV 558

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 2097797 seqs, 463214858 residues

Word size : 8

Total number of hits satisfying chosen parameters: 2975

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

Published Applications AA Main:*

- 1: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
- 2: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
- 3: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
- 4: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
- 5: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
- 6: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	558	100.0	558	6	US-11-134-563-12
2	558	100.0	558	6	US-11-052-554A-20
3	13	2.3	166	5	US-10-470-048B-73
4	12	2.2	107	4	US-10-424-599-207017
5	12	2.2	278	4	US-10-425-115-366802
6	12	2.2	361	4	US-10-017-161-1786
7	12	2.2	361	4	US-10-292-798-1452
8	12	2.2	367	4	US-10-424-599-144079
9	12	2.2	561	6	US-11-097-143-30426
10	12	2.2	785	5	US-10-741-849-7161
11	12	2.2	802	6	US-11-188-298-5674
12	12	2.2	802	6	US-11-188-298-20820
13	12	2.2	986	5	US-10-363-946-2
14	12	2.2	1026	6	US-11-097-143-19809
15	12	2.2	1026	6	US-11-097-143-38604
16	12	2.2	1695	6	US-11-097-143-28662
17	12	2.2	1711	6	US-11-097-143-12270
18	12	2.2	2159	5	US-10-732-923-9917
19	12	2.2	3712	4	US-10-108-605-103
20	12	2.2	3712	4	US-10-037-417-48
21	12	2.2	3712	4	US-10-037-417-51
22	12	2.2	3712	6	US-10-037-143-21654
23	12	2.2	3712	6	US-11-019-711-48
24	12	2.2	3712	6	US-11-019-711-51
25	12	2.2	3913	4	US-10-334-143-45
26	12	2.2	4274	5	US-10-450-763-31331
27	12	2.2	4377	5	US-10-756-149-4917

28	12	2.2	4384	5	US-10-821-234-1120	Sequence 1120, Ap
29	12	2.2	4386	5	US-10-450-763-37734	Sequence 37734, A
30	12	2.2	4397	5	US-10-450-763-52303	Sequence 52303, A
31	11	2.0	12	6	US-11-103-356A-3	Sequence 5, Appl
32	11	2.0	12	6	US-11-103-356A-7	Sequence 7, Appl
33	11	2.0	13	4	US-10-029-212-10	Sequence 10, Appl
34	11	2.0	24	3	US-09-864-761-40540	Sequence 40540, A
35	11	2.0	24	3	US-09-864-761-41883	Sequence 41883, A
36	11	2.0	39	5	US-10-622-893A-10	Sequence 10, Appl
37	11	2.0	53	4	US-10-425-115-213925	Sequence 213925, A
38	11	2.0	57	2	US-08-900-230-59	Sequence 59, Appl
39	11	2.0	65	4	US-10-437-963-163390	Sequence 163390, A
40	11	2.0	72	6	US-11-097-143-37041	Sequence 37041, A
41	11	2.0	75	4	US-10-425-115-355830	Sequence 355830, A
42	11	2.0	86	4	US-10-424-599-273144	Sequence 273144, A
43	11	2.0	86	4	US-10-424-599-161311	Sequence 161311, A
44	11	2.0	128	5	US-10-450-763-36460	Sequence 36460, A
45	11	2.0	141	4	US-10-424-599-201543	Sequence 201543, A
46	11	2.0	141	4	US-10-437-963-130292	Sequence 130292, A
47	11	2.0	147	5	US-10-739-930-8780	Sequence 8780, Ap
48	11	2.0	147	6	US-11-188-298-3226	Sequence 3226, Ap
49	11	2.0	157	6	US-11-188-298-7982	Sequence 7982, Ap
50	11	2.0	172	6	US-11-097-143-36390	Sequence 36390, A
51	11	2.0	182	6	US-11-097-143-38193	Sequence 38193, A
52	11	2.0	183	4	US-10-767-701-59312	Sequence 59312, A
53	11	2.0	198	4	US-10-424-599-159692	Sequence 159692, A
54	11	2.0	200	4	US-10-424-599-174641	Sequence 174641, A
55	11	2.0	203	6	US-11-097-143-38034	Sequence 38034, A
56	11	2.0	229	3	US-09-984-130-46	Sequence 46, Appl
57	11	2.0	229	3	US-09-836-353A-46	Sequence 46, Appl
58	11	2.0	229	6	US-11-097-143-35514	Sequence 35514, A
59	11	2.0	233	5	US-10-450-763-36461	Sequence 36461, A
60	11	2.0	237	4	US-10-437-963-189554	Sequence 189554, A
61	11	2.0	239	6	US-11-097-143-37722	Sequence 37722, A
62	11	2.0	242	6	US-11-097-143-39639	Sequence 39639, A
63	11	2.0	255	6	US-11-097-143-31008	Sequence 31008, A
64	11	2.0	267	4	US-10-425-114-54605	Sequence 54605, A
65	11	2.0	288	3	US-09-216-393-341	Sequence 341, Appl
66	11	2.0	288	3	US-09-216-393-344	Sequence 344, Appl
67	11	2.0	288	4	US-10-321-856-341	Sequence 341, Appl
68	11	2.0	289	4	US-10-321-856-344	Sequence 344, Appl
69	11	2.0	289	5	US-10-732-923-5938	Sequence 5938, Ap
70	11	2.0	291	4	US-10-767-701-41533	Sequence 41533, A
71	11	2.0	308	4	US-10-369-493-13774	Sequence 13774, A
72	11	2.0	311	4	US-10-369-493-4200	Sequence 4200, Ap
73	11	2.0	344	5	US-10-732-923-8440	Sequence 8440, Ap
74	11	2.0	363	4	US-10-437-963-119055	Sequence 119055, A
75	11	2.0	364	3	US-09-984-130-39	Sequence 39, Appl
76	11	2.0	364	3	US-09-836-353A-39	Sequence 39, Appl
77	11	2.0	390	6	US-11-098-688-10913	Sequence 10913, A
78	11	2.0	393	5	US-10-417-375-145	Sequence 145, Appl
79	11	2.0	422	4	US-10-032-588-7873	Sequence 7873, Ap
80	11	2.0	423	3	US-09-778-510-22	Sequence 22, Appl
81	11	2.0	423	3	US-09-778-187B-4	Sequence 4, Appl
82	11	2.0	423	4	US-10-302-021-22	Sequence 22, Appl
83	11	2.0	423	4	US-10-622-237-4	Sequence 4, Appl
84	11	2.0	423	5	US-10-898-408-4	Sequence 4, Appl
85	11	2.0	433	6	US-11-097-143-15147	Sequence 15147, A
86	11	2.0	435	4	US-10-032-588-7260	Sequence 7260, Ap
87	11	2.0	439	5	US-10-287-443A-249	Sequence 249, Appl
88	11	2.0	440	3	US-09-866-028-61	Sequence 61, Appl
89	11	2.0	440	3	US-09-944-443-61	Sequence 61, Appl
90	11	2.0	440	3	US-09-944-443-61	Sequence 61, Appl
91	11	2.0	440	3	US-09-944-443-61	Sequence 61, Appl
92	11	2.0	440	3	US-09-944-443-61	Sequence 61, Appl
93	11	2.0	440	3	US-09-944-443-61	Sequence 61, Appl
94	11	2.0	440	3	US-09-944-443-61	Sequence 61, Appl
95	11	2.0	440	3	US-09-944-443-61	Sequence 61, Appl
96	11	2.0	440	3	US-09-944-443-61	Sequence 61, Appl
97	11	2.0	440	3	US-09-944-443-61	Sequence 61, Appl
98	11	2.0	440	3	US-09-944-443-61	Sequence 61, Appl
99	11	2.0	440	3	US-09-944-443-61	Sequence 61, Appl
100	11	2.0	440	3	US-09-944-443-61	Sequence 61, Appl

101	11	2.0	440	3	US-09-944-896-61	Sequence 61, Appl	174	11	2.0	440	4	US-10-184-635-34	Sequence 34, Appl
102	11	2.0	440	3	US-09-944-944-61	Sequence 61, Appl	175	11	2.0	440	4	US-10-184-637-34	Sequence 34, Appl
103	11	2.0	440	3	US-09-944-929-61	Sequence 61, Appl	176	11	2.0	440	4	US-10-184-646-34	Sequence 34, Appl
104	11	2.0	440	3	US-09-944-907-61	Sequence 61, Appl	177	11	2.0	440	4	US-10-184-647-34	Sequence 34, Appl
105	11	2.0	440	3	US-09-944-884-61	Sequence 61, Appl	178	11	2.0	440	4	US-10-184-652-34	Sequence 34, Appl
106	11	2.0	440	3	US-09-944-852-61	Sequence 61, Appl	179	11	2.0	440	4	US-10-187-594-34	Sequence 34, Appl
107	11	2.0	440	3	US-09-943-780-61	Sequence 61, Appl	180	11	2.0	440	4	US-10-187-596-34	Sequence 34, Appl
108	11	2.0	440	3	US-09-945-584-61	Sequence 61, Appl	181	11	2.0	440	4	US-10-187-745-34	Sequence 34, Appl
109	11	2.0	440	3	US-09-943-686-61	Sequence 61, Appl	182	11	2.0	440	4	US-10-187-885-34	Sequence 34, Appl
110	11	2.0	440	3	US-09-943-664-61	Sequence 61, Appl	183	11	2.0	440	4	US-10-187-886-34	Sequence 34, Appl
111	11	2.0	440	4	US-10-174-590-34	Sequence 34, Appl	184	11	2.0	440	4	US-10-199-664-34	Sequence 34, Appl
112	11	2.0	440	4	US-10-176-758-34	Sequence 34, Appl	185	11	2.0	440	4	US-10-196-756-34	Sequence 34, Appl
113	11	2.0	440	4	US-10-175-737-34	Sequence 34, Appl	186	11	2.0	440	4	US-10-176-751-34	Sequence 34, Appl
114	11	2.0	440	4	US-10-174-581-34	Sequence 34, Appl	187	11	2.0	440	4	US-10-176-760-34	Sequence 34, Appl
115	11	2.0	440	4	US-10-176-483-34	Sequence 34, Appl	188	11	2.0	440	4	US-10-176-990-34	Sequence 34, Appl
116	11	2.0	440	4	US-10-176-749-34	Sequence 34, Appl	189	11	2.0	440	4	US-10-180-541-34	Sequence 34, Appl
117	11	2.0	440	4	US-10-176-914-34	Sequence 34, Appl	190	11	2.0	440	4	US-10-180-542-34	Sequence 34, Appl
118	11	2.0	440	4	US-10-176-915-34	Sequence 34, Appl	191	11	2.0	440	4	US-10-180-548-34	Sequence 34, Appl
119	11	2.0	440	4	US-10-173-706-34	Sequence 34, Appl	192	11	2.0	440	4	US-10-180-551-34	Sequence 34, Appl
120	11	2.0	440	4	US-10-175-738-34	Sequence 34, Appl	193	11	2.0	440	4	US-10-180-998-34	Sequence 34, Appl
121	11	2.0	440	4	US-10-175-752-34	Sequence 34, Appl	194	11	2.0	440	4	US-10-180-999-34	Sequence 34, Appl
122	11	2.0	440	4	US-10-176-482-34	Sequence 34, Appl	195	11	2.0	440	4	US-10-183-013-34	Sequence 34, Appl
123	11	2.0	440	4	US-10-176-757-34	Sequence 34, Appl	196	11	2.0	440	4	US-10-184-612-34	Sequence 34, Appl
124	11	2.0	440	4	US-10-176-913-34	Sequence 34, Appl	197	11	2.0	440	4	US-10-184-616-34	Sequence 34, Appl
125	11	2.0	440	4	US-10-180-552-34	Sequence 34, Appl	198	11	2.0	440	4	US-10-184-617-34	Sequence 34, Appl
126	11	2.0	440	4	US-10-180-557-34	Sequence 34, Appl	199	11	2.0	440	4	US-10-184-622-34	Sequence 34, Appl
127	11	2.0	440	4	US-10-173-700-34	Sequence 34, Appl	200	11	2.0	440	4	US-10-184-628-34	Sequence 34, Appl
128	11	2.0	440	4	US-10-174-572-34	Sequence 34, Appl	201	11	2.0	440	4	US-10-184-629-34	Sequence 34, Appl
129	11	2.0	440	4	US-10-174-579-34	Sequence 34, Appl	202	11	2.0	440	4	US-10-184-630-34	Sequence 34, Appl
130	11	2.0	440	4	US-10-174-582-34	Sequence 34, Appl	203	11	2.0	440	4	US-10-184-631-34	Sequence 34, Appl
131	11	2.0	440	4	US-10-174-588-34	Sequence 34, Appl	204	11	2.0	440	4	US-10-184-632-34	Sequence 34, Appl
132	11	2.0	440	4	US-10-175-739-34	Sequence 34, Appl	205	11	2.0	440	4	US-10-184-636-34	Sequence 34, Appl
133	11	2.0	440	4	US-10-175-740-34	Sequence 34, Appl	206	11	2.0	440	4	US-10-184-640-34	Sequence 34, Appl
134	11	2.0	440	4	US-10-175-743-34	Sequence 34, Appl	207	11	2.0	440	4	US-10-184-650-34	Sequence 34, Appl
135	11	2.0	440	4	US-10-176-488-34	Sequence 34, Appl	208	11	2.0	440	4	US-10-184-651-34	Sequence 34, Appl
136	11	2.0	440	4	US-10-176-492-34	Sequence 34, Appl	209	11	2.0	440	4	US-10-187-588-34	Sequence 34, Appl
137	11	2.0	440	4	US-10-176-747-34	Sequence 34, Appl	210	11	2.0	440	4	US-10-187-597-34	Sequence 34, Appl
138	11	2.0	440	4	US-10-176-750-34	Sequence 34, Appl	211	11	2.0	440	4	US-10-187-598-34	Sequence 34, Appl
139	11	2.0	440	4	US-10-176-985-34	Sequence 34, Appl	212	11	2.0	440	4	US-10-187-600-34	Sequence 34, Appl
140	11	2.0	440	4	US-10-176-987-34	Sequence 34, Appl	213	11	2.0	440	4	US-10-187-601-34	Sequence 34, Appl
141	11	2.0	440	4	US-10-176-992-34	Sequence 34, Appl	214	11	2.0	440	4	US-10-187-602-34	Sequence 34, Appl
142	11	2.0	440	4	US-10-176-993-34	Sequence 34, Appl	215	11	2.0	440	4	US-10-187-603-34	Sequence 34, Appl
143	11	2.0	440	4	US-10-184-658-34	Sequence 34, Appl	216	11	2.0	440	4	US-10-187-741-34	Sequence 34, Appl
144	11	2.0	440	4	US-10-176-991-34	Sequence 34, Appl	217	11	2.0	440	4	US-10-187-743-34	Sequence 34, Appl
145	11	2.0	440	4	US-10-173-695-34	Sequence 34, Appl	218	11	2.0	440	4	US-10-187-746-34	Sequence 34, Appl
146	11	2.0	440	4	US-10-173-697-34	Sequence 34, Appl	219	11	2.0	440	4	US-10-187-747-34	Sequence 34, Appl
147	11	2.0	440	4	US-10-173-705-34	Sequence 34, Appl	220	11	2.0	440	4	US-10-187-751-34	Sequence 34, Appl
148	11	2.0	440	4	US-10-174-576-34	Sequence 34, Appl	221	11	2.0	440	4	US-10-187-753-34	Sequence 34, Appl
149	11	2.0	440	4	US-10-174-585-34	Sequence 34, Appl	222	11	2.0	440	4	US-10-187-754-34	Sequence 34, Appl
150	11	2.0	440	4	US-10-174-586-34	Sequence 34, Appl	223	11	2.0	440	4	US-10-187-757-34	Sequence 34, Appl
151	11	2.0	440	4	US-10-175-747-34	Sequence 34, Appl	224	11	2.0	440	4	US-10-187-884-34	Sequence 34, Appl
152	11	2.0	440	4	US-10-176-481-34	Sequence 34, Appl	225	11	2.0	440	4	US-10-188-767-34	Sequence 34, Appl
153	11	2.0	440	4	US-10-176-485-34	Sequence 34, Appl	226	11	2.0	440	4	US-10-188-769-34	Sequence 34, Appl
154	11	2.0	440	4	US-10-176-487-34	Sequence 34, Appl	227	11	2.0	440	4	US-10-188-770-34	Sequence 34, Appl
155	11	2.0	440	4	US-10-176-493-34	Sequence 34, Appl	228	11	2.0	440	4	US-10-188-773-34	Sequence 34, Appl
156	11	2.0	440	4	US-10-176-756-34	Sequence 34, Appl	229	11	2.0	440	4	US-10-188-781-34	Sequence 34, Appl
157	11	2.0	440	4	US-10-176-911-34	Sequence 34, Appl	230	11	2.0	440	4	US-10-194-361-34	Sequence 34, Appl
158	11	2.0	440	4	US-10-176-919-34	Sequence 34, Appl	231	11	2.0	440	4	US-10-194-423-34	Sequence 34, Appl
159	11	2.0	440	4	US-10-176-925-34	Sequence 34, Appl	232	11	2.0	440	4	US-10-195-897-34	Sequence 34, Appl
160	11	2.0	440	4	US-10-176-978-34	Sequence 34, Appl	233	11	2.0	440	4	US-10-195-901-34	Sequence 34, Appl
161	11	2.0	440	4	US-10-179-510-34	Sequence 34, Appl	234	11	2.0	440	4	US-10-195-902-34	Sequence 34, Appl
162	11	2.0	440	4	US-10-180-543-34	Sequence 34, Appl	235	11	2.0	440	4	US-10-196-743-34	Sequence 34, Appl
163	11	2.0	440	4	US-10-180-544-34	Sequence 34, Appl	236	11	2.0	440	4	US-10-196-760-34	Sequence 34, Appl
164	11	2.0	440	4	US-10-180-546-34	Sequence 34, Appl	237	11	2.0	440	4	US-10-193-708-34	Sequence 34, Appl
165	11	2.0	440	4	US-10-180-547-34	Sequence 34, Appl	238	11	2.0	440	4	US-10-176-479-34	Sequence 34, Appl
166	11	2.0	440	4	US-10-180-549-34	Sequence 34, Appl	239	11	2.0	440	4	US-10-176-748-34	Sequence 34, Appl
167	11	2.0	440	4	US-10-180-555-34	Sequence 34, Appl	240	11	2.0	440	4	US-10-176-916-34	Sequence 34, Appl
168	11	2.0	440	4	US-10-180-559-34	Sequence 34, Appl	241	11	2.0	440	4	US-10-179-507-34	Sequence 34, Appl
169	11	2.0	440	4	US-10-181-000-34	Sequence 34, Appl	242	11	2.0	440	4	US-10-179-516-34	Sequence 34, Appl
170	11	2.0	440	4	US-10-183-010-34	Sequence 34, Appl	243	11	2.0	440	4	US-10-179-519-34	Sequence 34, Appl
171	11	2.0	440	4	US-10-183-012-34	Sequence 34, Appl	244	11	2.0	440	4	US-10-179-525-34	Sequence 34, Appl
172	11	2.0	440	4	US-10-184-614-34	Sequence 34, Appl	245	11	2.0	440	4	US-10-180-540-34	Sequence 34, Appl
173	11	2.0	440	4	US-10-184-623-34	Sequence 34, Appl	246	11	2.0	440	4	US-10-180-545-34	Sequence 34, Appl

393	11	2.0	440	4	US-10-206-928-34	Sequence 34, Appl	466	11	2.0	440	4	US-10-205-894-34	Sequence 34, Appl
394	11	2.0	440	4	US-10-207-914-34	Sequence 34, Appl	467	11	2.0	440	4	US-10-205-898-34	Sequence 34, Appl
395	11	2.0	440	4	US-10-207-921-34	Sequence 34, Appl	468	11	2.0	440	4	US-10-205-898-34	Sequence 34, Appl
396	11	2.0	440	4	US-10-207-922-34	Sequence 34, Appl	469	11	2.0	440	4	US-10-205-901-34	Sequence 34, Appl
397	11	2.0	440	4	US-10-208-027-34	Sequence 34, Appl	470	11	2.0	440	4	US-10-205-903-34	Sequence 34, Appl
398	11	2.0	440	4	US-10-196-757-34	Sequence 34, Appl	471	11	2.0	440	4	US-10-206-909-34	Sequence 34, Appl
399	11	2.0	440	4	US-10-196-754-34	Sequence 34, Appl	472	11	2.0	440	4	US-10-206-910-34	Sequence 34, Appl
400	11	2.0	440	4	US-10-174-571-34	Sequence 34, Appl	473	11	2.0	440	4	US-10-206-911-34	Sequence 34, Appl
401	11	2.0	440	4	US-10-176-746-34	Sequence 34, Appl	474	11	2.0	440	4	US-10-206-912-34	Sequence 34, Appl
402	11	2.0	440	4	US-10-176-923-34	Sequence 34, Appl	475	11	2.0	440	4	US-10-206-913-34	Sequence 34, Appl
403	11	2.0	440	4	US-10-183-011-34	Sequence 34, Appl	476	11	2.0	440	4	US-10-206-914-34	Sequence 34, Appl
404	11	2.0	440	4	US-10-184-633-34	Sequence 34, Appl	477	11	2.0	440	4	US-10-206-920-34	Sequence 34, Appl
405	11	2.0	440	4	US-10-184-639-34	Sequence 34, Appl	478	11	2.0	440	4	US-10-206-921-34	Sequence 34, Appl
406	11	2.0	440	4	US-10-187-742-34	Sequence 34, Appl	479	11	2.0	440	4	US-10-206-923-34	Sequence 34, Appl
407	11	2.0	440	4	US-10-187-748-34	Sequence 34, Appl	480	11	2.0	440	4	US-10-206-925-34	Sequence 34, Appl
408	11	2.0	440	4	US-10-188-766-34	Sequence 34, Appl	481	11	2.0	440	4	US-10-206-926-34	Sequence 34, Appl
409	11	2.0	440	4	US-10-188-771-34	Sequence 34, Appl	482	11	2.0	440	4	US-10-206-927-34	Sequence 34, Appl
410	11	2.0	440	4	US-10-192-006-34	Sequence 34, Appl	483	11	2.0	440	4	US-10-207-916-34	Sequence 34, Appl
411	11	2.0	440	4	US-10-192-008-34	Sequence 34, Appl	484	11	2.0	440	4	US-10-207-917-34	Sequence 34, Appl
412	11	2.0	440	4	US-10-192-009-34	Sequence 34, Appl	485	11	2.0	440	4	US-10-207-918-34	Sequence 34, Appl
413	11	2.0	440	4	US-10-192-012-34	Sequence 34, Appl	486	11	2.0	440	4	US-10-207-919-34	Sequence 34, Appl
414	11	2.0	440	4	US-10-192-014-34	Sequence 34, Appl	487	11	2.0	440	4	US-10-207-920-34	Sequence 34, Appl
415	11	2.0	440	4	US-10-192-016-34	Sequence 34, Appl	488	11	2.0	440	4	US-10-208-021-34	Sequence 34, Appl
416	11	2.0	440	4	US-10-194-362-34	Sequence 34, Appl	489	11	2.0	440	4	US-10-208-022-34	Sequence 34, Appl
417	11	2.0	440	4	US-10-194-364-34	Sequence 34, Appl	490	11	2.0	440	4	US-10-208-023-34	Sequence 34, Appl
418	11	2.0	440	4	US-10-194-395-34	Sequence 34, Appl	491	11	2.0	440	4	US-10-208-026-34	Sequence 34, Appl
419	11	2.0	440	4	US-10-194-424-34	Sequence 34, Appl	492	11	2.0	440	4	US-10-208-029-34	Sequence 34, Appl
420	11	2.0	440	4	US-10-194-458-34	Sequence 34, Appl	493	11	2.0	440	4	US-10-208-029-34	Sequence 34, Appl
421	11	2.0	440	4	US-10-194-459-34	Sequence 34, Appl	494	11	2.0	440	4	US-10-208-030-34	Sequence 34, Appl
422	11	2.0	440	4	US-10-194-488-34	Sequence 34, Appl	495	11	2.0	440	4	US-10-232-232-34	Sequence 34, Appl
423	11	2.0	440	4	US-10-195-886-34	Sequence 34, Appl	496	11	2.0	440	4	US-10-135-898-34	Sequence 34, Appl
424	11	2.0	440	4	US-10-195-891-34	Sequence 34, Appl	497	11	2.0	440	4	US-10-136-759-34	Sequence 34, Appl
425	11	2.0	440	4	US-10-196-746-34	Sequence 34, Appl	498	11	2.0	440	4	US-10-173-693-34	Sequence 34, Appl
426	11	2.0	440	4	US-10-196-752-34	Sequence 34, Appl	499	11	2.0	440	4	US-10-174-578-34	Sequence 34, Appl
427	11	2.0	440	4	US-10-196-753-34	Sequence 34, Appl	500	11	2.0	440	4	US-10-175-741-34	Sequence 34, Appl
428	11	2.0	440	4	US-10-196-761-34	Sequence 34, Appl	501	11	2.0	440	4	US-10-175-750-34	Sequence 34, Appl
429	11	2.0	440	4	US-10-197-692-34	Sequence 34, Appl	502	11	2.0	440	4	US-10-176-966-34	Sequence 34, Appl
430	11	2.0	440	4	US-10-197-693-34	Sequence 34, Appl	503	11	2.0	440	4	US-10-184-641-34	Sequence 34, Appl
431	11	2.0	440	4	US-10-197-696-34	Sequence 34, Appl	504	11	2.0	440	4	US-10-187-888-34	Sequence 34, Appl
432	11	2.0	440	4	US-10-197-698-34	Sequence 34, Appl	505	11	2.0	440	4	US-10-194-360-34	Sequence 34, Appl
433	11	2.0	440	4	US-10-197-703-34	Sequence 34, Appl	506	11	2.0	440	4	US-10-194-365-34	Sequence 34, Appl
434	11	2.0	440	4	US-10-197-711-34	Sequence 34, Appl	507	11	2.0	440	4	US-10-195-895-34	Sequence 34, Appl
435	11	2.0	440	4	US-10-198-757-34	Sequence 34, Appl	508	11	2.0	440	4	US-10-201-323-34	Sequence 34, Appl
436	11	2.0	440	4	US-10-198-761-34	Sequence 34, Appl	509	11	2.0	440	4	US-10-201-323-34	Sequence 34, Appl
437	11	2.0	440	4	US-10-198-762-34	Sequence 34, Appl	510	11	2.0	440	4	US-10-205-510-34	Sequence 34, Appl
438	11	2.0	440	4	US-10-198-763-34	Sequence 34, Appl	511	11	2.0	440	4	US-10-205-891-34	Sequence 34, Appl
439	11	2.0	440	4	US-10-198-767-34	Sequence 34, Appl	512	11	2.0	440	4	US-10-206-917-34	Sequence 34, Appl
440	11	2.0	440	4	US-10-199-301-34	Sequence 34, Appl	513	11	2.0	440	4	US-10-207-923-34	Sequence 34, Appl
441	11	2.0	440	4	US-10-199-307-34	Sequence 34, Appl	514	11	2.0	440	4	US-10-207-924-34	Sequence 34, Appl
442	11	2.0	440	4	US-10-199-312-34	Sequence 34, Appl	515	11	2.0	440	4	US-10-208-028-34	Sequence 34, Appl
443	11	2.0	440	4	US-10-199-315-34	Sequence 34, Appl	516	11	2.0	440	4	US-10-205-904-34	Sequence 34, Appl
444	11	2.0	440	4	US-10-199-316-34	Sequence 34, Appl	517	11	2.0	440	4	US-10-175-753-34	Sequence 34, Appl
445	11	2.0	440	4	US-10-199-457-34	Sequence 34, Appl	518	11	2.0	440	4	US-10-180-553-34	Sequence 34, Appl
446	11	2.0	440	4	US-10-199-459-34	Sequence 34, Appl	519	11	2.0	440	4	US-10-201-327-34	Sequence 34, Appl
447	11	2.0	440	4	US-10-199-460-34	Sequence 34, Appl	520	11	2.0	440	4	US-10-121-062-34	Sequence 34, Appl
448	11	2.0	440	4	US-10-199-461-34	Sequence 34, Appl	521	11	2.0	440	4	US-10-183-003-34	Sequence 34, Appl
449	11	2.0	440	4	US-10-199-667-34	Sequence 34, Appl	522	11	2.0	440	4	US-10-187-691-34	Sequence 34, Appl
450	11	2.0	440	4	US-10-199-673-34	Sequence 34, Appl	523	11	2.0	440	4	US-10-177-696-34	Sequence 34, Appl
451	11	2.0	440	4	US-10-201-321-34	Sequence 34, Appl	524	11	2.0	440	4	US-10-175-575A-34	Sequence 34, Appl
452	11	2.0	440	4	US-10-201-322-34	Sequence 34, Appl	525	11	2.0	440	4	US-10-176-979-34	Sequence 34, Appl
453	11	2.0	440	4	US-10-201-326-34	Sequence 34, Appl	526	11	2.0	440	4	US-10-187-592-34	Sequence 34, Appl
454	11	2.0	440	4	US-10-201-532-34	Sequence 34, Appl	527	11	2.0	440	4	US-10-187-592-34	Sequence 34, Appl
455	11	2.0	440	4	US-10-201-533-34	Sequence 34, Appl	528	11	2.0	440	4	US-10-197-691-34	Sequence 34, Appl
456	11	2.0	440	4	US-10-201-535-34	Sequence 34, Appl	529	11	2.0	440	4	US-10-198-771-34	Sequence 34, Appl
457	11	2.0	440	4	US-10-201-769-34	Sequence 34, Appl	530	11	2.0	440	4	US-10-174-575A-34	Sequence 34, Appl
458	11	2.0	440	4	US-10-201-771-34	Sequence 34, Appl	531	11	2.0	440	4	US-10-179-520-34	Sequence 34, Appl
459	11	2.0	440	4	US-10-201-854-34	Sequence 34, Appl	532	11	2.0	440	4	US-10-201-325-34	Sequence 34, Appl
460	11	2.0	440	4	US-10-202-410-34	Sequence 34, Appl	533	11	2.0	440	4	US-10-202-941-34	Sequence 34, Appl
461	11	2.0	440	4	US-10-202-473-34	Sequence 34, Appl	534	11	2.0	440	4	US-10-208-910-34	Sequence 34, Appl
462	11	2.0	440	4	US-10-202-474-34	Sequence 34, Appl	535	11	2.0	440	4	US-10-179-526-34	Sequence 34, Appl
463	11	2.0	440	4	US-10-205-503-34	Sequence 34, Appl	536	11	2.0	440	4	US-10-173-701-34	Sequence 34, Appl
464	11	2.0	440	4	US-10-205-512-34	Sequence 34, Appl	537	11	2.0	440	4	US-10-179-511-34	Sequence 34, Appl
465	11	2.0	440	4	US-10-205-892-34	Sequence 34, Appl	538	11	2.0	440	4	US-10-179-518-34	Sequence 34, Appl

539	11	2.0	440	4	US-10-183-018-34	Sequence 34, Appl	612	11	2.0	440	4	US-10-194-486-34	Sequence 34, Appl
540	11	2.0	440	4	US-10-184-624-34	Sequence 34, Appl	613	11	2.0	440	4	US-10-195-800-34	Sequence 34, Appl
541	11	2.0	440	4	US-10-184-657-34	Sequence 34, Appl	614	11	2.0	440	4	US-10-198-759-34	Sequence 34, Appl
542	11	2.0	440	4	US-10-197-701-34	Sequence 34, Appl	615	11	2.0	440	4	US-10-205-506-34	Sequence 34, Appl
543	11	2.0	440	4	US-10-197-706-34	Sequence 34, Appl	616	11	2.0	440	4	US-10-429-667-61	Sequence 61, Appl
544	11	2.0	440	4	US-10-201-857-34	Sequence 34, Appl	617	11	2.0	440	4	US-10-174-570-34	Sequence 34, Appl
545	11	2.0	440	4	US-10-202-413-34	Sequence 34, Appl	618	11	2.0	440	4	US-10-183-005-34	Sequence 34, Appl
546	11	2.0	440	4	US-10-202-938-34	Sequence 34, Appl	619	11	2.0	440	4	US-10-179-523-34	Sequence 34, Appl
547	11	2.0	440	4	US-10-202-940-34	Sequence 34, Appl	620	11	2.0	440	4	US-10-199-463-34	Sequence 34, Appl
548	11	2.0	440	4	US-10-205-508-34	Sequence 34, Appl	621	11	2.0	440	4	US-10-202-471-34	Sequence 34, Appl
549	11	2.0	440	4	US-10-205-505-34	Sequence 34, Appl	622	11	2.0	440	4	US-10-207-915-34	Sequence 34, Appl
550	11	2.0	440	4	US-10-206-918-34	Sequence 34, Appl	623	11	2.0	440	4	US-10-197-709-34	Sequence 34, Appl
551	11	2.0	440	4	US-10-208-025-34	Sequence 34, Appl	624	11	2.0	440	4	US-10-206-915-34	Sequence 34, Appl
552	11	2.0	440	4	US-10-198-760-34	Sequence 34, Appl	625	11	2.0	440	4	US-10-199-670-34	Sequence 34, Appl
553	11	2.0	440	4	US-10-201-772-34	Sequence 34, Appl	626	11	2.0	440	4	US-10-201-858-34	Sequence 34, Appl
554	11	2.0	440	4	US-10-184-613-34	Sequence 34, Appl	627	11	2.0	440	4	US-10-205-890-34	Sequence 34, Appl
555	11	2.0	440	4	US-10-187-739-34	Sequence 34, Appl	628	11	2.0	440	4	US-10-208-024-34	Sequence 34, Appl
556	11	2.0	440	4	US-10-206-907-34	Sequence 34, Appl	629	11	2.0	440	4	US-10-201-853-34	Sequence 34, Appl
557	11	2.0	440	4	US-10-183-009-34	Sequence 34, Appl	630	11	2.0	440	4	US-10-677-471-61	Sequence 61, Appl
558	11	2.0	440	4	US-10-187-755-34	Sequence 34, Appl	631	11	2.0	440	4	US-10-677-669-61	Sequence 61, Appl
559	11	2.0	440	4	US-10-199-672-34	Sequence 34, Appl	632	11	2.0	440	4	US-10-206-916-34	Sequence 61, Appl
560	11	2.0	440	4	US-10-187-749-34	Sequence 34, Appl	633	11	2.0	440	5	US-10-735-014-61	Sequence 61, Appl
561	11	2.0	440	4	US-10-194-457-34	Sequence 34, Appl	634	11	2.0	440	5	US-10-854-947-61	Sequence 61, Appl
562	11	2.0	440	4	US-10-184-642-34	Sequence 34, Appl	635	11	2.0	440	5	US-10-858-993-61	Sequence 61, Appl
563	11	2.0	440	4	US-10-196-747-34	Sequence 34, Appl	636	11	2.0	440	5	US-10-901-400-61	Sequence 61, Appl
564	11	2.0	440	4	US-10-173-689-34	Sequence 34, Appl	637	11	2.0	440	5	US-10-858-981-61	Sequence 61, Appl
565	11	2.0	440	4	US-10-173-690-34	Sequence 34, Appl	638	11	2.0	440	5	US-10-899-671-61	Sequence 61, Appl
566	11	2.0	440	4	US-10-173-691-34	Sequence 34, Appl	639	11	2.0	440	5	US-10-943-353-61	Sequence 61, Appl
567	11	2.0	440	4	US-10-173-694-34	Sequence 34, Appl	640	11	2.0	440	5	US-10-183-001-34	Sequence 34, Appl
568	11	2.0	440	4	US-10-173-698-34	Sequence 34, Appl	641	11	2.0	440	5	US-10-175-749-34	Sequence 34, Appl
569	11	2.0	440	4	US-10-173-699-34	Sequence 34, Appl	642	11	2.0	440	5	US-10-180-555-34	Sequence 34, Appl
570	11	2.0	440	4	US-10-173-707-34	Sequence 34, Appl	643	11	2.0	440	5	US-10-195-883-34	Sequence 34, Appl
571	11	2.0	440	4	US-10-174-569-34	Sequence 34, Appl	644	11	2.0	440	5	US-10-195-888-34	Sequence 34, Appl
572	11	2.0	440	4	US-10-174-583-34	Sequence 34, Appl	645	11	2.0	440	5	US-10-194-487-34	Sequence 34, Appl
573	11	2.0	440	4	US-10-174-587-34	Sequence 34, Appl	646	11	2.0	440	5	US-10-195-889-34	Sequence 34, Appl
574	11	2.0	440	4	US-10-174-589-34	Sequence 34, Appl	647	11	2.0	442	3	US-09-778-510-20	Sequence 20, Appl
575	11	2.0	440	4	US-10-174-591-34	Sequence 34, Appl	648	11	2.0	442	3	US-09-984-130-136	Sequence 136, Appl
576	11	2.0	440	4	US-10-175-736-34	Sequence 34, Appl	649	11	2.0	442	3	US-09-836-353A-136	Sequence 20, Appl
577	11	2.0	440	4	US-10-175-742-34	Sequence 34, Appl	650	11	2.0	442	4	US-10-302-041-20	Sequence 20, Appl
578	11	2.0	440	4	US-10-175-744-34	Sequence 34, Appl	651	11	2.0	442	4	US-10-015-115-110	Sequence 110, Appl
579	11	2.0	440	4	US-10-175-745-34	Sequence 34, Appl	652	11	2.0	442	4	US-10-015-115-111	Sequence 111, Appl
580	11	2.0	440	4	US-10-175-748-34	Sequence 34, Appl	653	11	2.0	442	4	US-10-363-616-262	Sequence 262, Appl
581	11	2.0	440	4	US-10-175-751-34	Sequence 34, Appl	654	11	2.0	442	4	US-10-622-237-2	Sequence 2, Appl
582	11	2.0	440	4	US-10-175-754-34	Sequence 34, Appl	655	11	2.0	442	4	US-10-898-408-2	Sequence 13422, A
583	11	2.0	440	4	US-10-176-480-34	Sequence 34, Appl	656	11	2.0	442	4	US-11-097-143-13422	Sequence 112, App
584	11	2.0	440	4	US-10-176-489-34	Sequence 34, Appl	657	11	2.0	442	4	US-11-087-099-7775	Sequence 7775, Ap
585	11	2.0	440	4	US-10-176-754-34	Sequence 34, Appl	658	11	2.0	445	6	US-10-425-114-62647	Sequence 62647, A
586	11	2.0	440	4	US-10-176-755-34	Sequence 34, Appl	659	11	2.0	445	6	US-11-097-143-17016	Sequence 17016, A
587	11	2.0	440	4	US-10-176-759-34	Sequence 34, Appl	660	11	2.0	445	6	US-10-015-115-113	Sequence 113, App
588	11	2.0	440	4	US-10-176-920-34	Sequence 34, Appl	661	11	2.0	460	4	US-10-424-598-7589	Sequence 7589, Ap
589	11	2.0	440	4	US-10-176-922-34	Sequence 34, Appl	662	11	2.0	474	6	US-10-278-536-182	Sequence 182, App
590	11	2.0	440	4	US-10-176-924-34	Sequence 34, Appl	663	11	2.0	519	4	US-10-412-699B-618	Sequence 618, App
591	11	2.0	440	4	US-10-176-984-34	Sequence 34, Appl	664	11	2.0	520	6	US-11-097-143-16854	Sequence 16854, A
592	11	2.0	440	4	US-10-179-508-34	Sequence 34, Appl	665	11	2.0	523	4	US-10-017-161-1982	Sequence 1982, Ap
593	11	2.0	440	4	US-10-179-512-34	Sequence 34, Appl	666	11	2.0	523	4	US-10-292-799-1630	Sequence 1630, Ap
594	11	2.0	440	4	US-10-179-515-34	Sequence 34, Appl	667	11	2.0	523	4	US-10-425-114-58445	Sequence 58445, A
595	11	2.0	440	4	US-10-173-692-34	Sequence 34, Appl	668	11	2.0	524	4	US-10-437-961-177911	Sequence 177911, A
596	11	2.0	440	4	US-10-173-702-34	Sequence 34, Appl	669	11	2.0	524	4	US-10-425-115-360409	Sequence 360409, A
597	11	2.0	440	4	US-10-173-703-34	Sequence 34, Appl	670	11	2.0	545	5	US-10-287-438A-271	Sequence 271, App
598	11	2.0	440	4	US-10-173-704-34	Sequence 34, Appl	671	11	2.0	559	3	US-09-900-708-2	Sequence 2, Appl
599	11	2.0	440	4	US-10-174-574-34	Sequence 34, Appl	672	11	2.0	566	4	US-10-672-074-8	Sequence 8, Appl
600	11	2.0	440	4	US-10-176-486-34	Sequence 34, Appl	673	11	2.0	571	6	US-11-041-553-5	Sequence 5, Appl
601	11	2.0	440	4	US-10-176-490-34	Sequence 34, Appl	674	11	2.0	620	5	US-10-732-923-13041	Sequence 13041, A
602	11	2.0	440	4	US-10-176-752-34	Sequence 34, Appl	675	11	2.0	671	6	US-11-188-298-9026	Sequence 9026, Ap
603	11	2.0	440	4	US-10-176-981-34	Sequence 34, Appl	676	11	2.0	680	6	US-09-881-736-6	Sequence 6, Appl
604	11	2.0	440	4	US-10-176-983-34	Sequence 34, Appl	677	11	2.0	681	11	US-10-029-180-22	Sequence 22, Appl
605	11	2.0	440	4	US-10-176-988-34	Sequence 34, Appl	678	11	2.0	694	5	US-10-952-045-22	Sequence 22, Appl
606	11	2.0	440	4	US-10-179-517-34	Sequence 34, Appl	679	11	2.0	707	4	US-10-195-144-31	Sequence 31, Appl
607	11	2.0	440	4	US-10-179-521-34	Sequence 34, Appl	680	11	2.0	707	4	US-10-345-072-31	Sequence 31, Appl
608	11	2.0	440	4	US-10-802-475-34	Sequence 34, Appl	681	11	2.0	707	4		
609	11	2.0	440	4	US-10-195-887-34	Sequence 34, Appl	682	11	2.0	707	4		
610	11	2.0	440	4	US-10-195-893-34	Sequence 34, Appl	683	11	2.0	707	4		
611	11	2.0	440	4	US-10-179-509-34	Sequence 34, Appl	684	11	2.0	707	4		

685	11	2.0	725	6	US-11-097-143-32844	Sequence 32844, A	758	11	2.0	2014	4	US-10-141-756-51	Sequence 51, Appl
686	11	2.0	730	5	US-10-450-763-57590	Sequence 57590, A	759	11	2.0	2014	4	US-10-141-759-51	Sequence 51, Appl
687	11	2.0	746	6	US-11-097-143-4395	Sequence 4395, Ap	760	11	2.0	2014	4	US-10-140-805-51	Sequence 51, Appl
688	11	2.0	769	4	US-10-032-585-7117	Sequence 7117, Ap	761	11	2.0	2014	4	US-10-140-864-51	Sequence 51, Appl
689	11	2.0	793	5	US-10-732-923-3091	Sequence 3091, Ap	762	11	2.0	2027	4	US-10-123-155-175	Sequence 175, Ap
690	11	2.0	793	6	US-11-041-553-8	Sequence 8, Appl1	763	11	2.0	2027	4	US-10-146-731-175	Sequence 175, Ap
691	11	2.0	826	4	US-10-184-644-247	Sequence 247, App	764	11	2.0	2027	4	US-10-140-472-175	Sequence 175, Ap
692	11	2.0	826	4	US-10-184-634-247	Sequence 247, App	765	11	2.0	2027	4	US-10-141-761-175	Sequence 175, Ap
693	11	2.0	945	2	US-08-965-272-2	Sequence 2, Appl1	766	11	2.0	2027	4	US-10-142-885-175	Sequence 175, App
694	11	2.0	945	2	US-10-121-882-2	Sequence 2, Appl1	767	11	2.0	2027	4	US-10-158-790-175	Sequence 175, App
695	11	2.0	961	5	US-10-732-923-9614	Sequence 9614, Ap	768	11	2.0	2027	4	US-10-137-871-175	Sequence 175, App
696	11	2.0	1002	4	US-10-654-416-4	Sequence 4, Appl1	769	11	2.0	2027	4	US-10-140-923-175	Sequence 175, App
697	11	2.0	1024	5	US-10-732-923-13550	Sequence 13550, A	770	11	2.0	2027	4	US-10-141-756-175	Sequence 175, App
698	11	2.0	1057	6	US-11-097-143-23112	Sequence 23112, A	771	11	2.0	2027	4	US-10-141-759-175	Sequence 175, App
699	11	2.0	1077	4	US-10-032-585-79311	Sequence 7931, Ap	772	11	2.0	2027	4	US-10-140-805-175	Sequence 175, App
700	11	2.0	1088	4	US-10-267-502-434	Sequence 434, App	773	11	2.0	2027	4	US-10-140-864-175	Sequence 175, App
701	11	2.0	1088	6	US-11-097-143-30609	Sequence 30609, A	774	11	2.0	2120	4	US-10-123-155-73	Sequence 73, Appl
702	11	2.0	1128	4	US-10-156-240-11	Sequence 11, Appl	775	11	2.0	2120	4	US-10-146-731-73	Sequence 73, Appl
703	11	2.0	1137	6	US-11-241-631-11	Sequence 11, Appl	776	11	2.0	2120	4	US-10-140-472-73	Sequence 73, Appl
704	11	2.0	1211	6	US-11-097-143-2400	Sequence 2400, Ap	777	11	2.0	2120	4	US-10-141-761-73	Sequence 73, Appl
705	11	2.0	1211	6	US-11-188-298-10688	Sequence 10688, A	778	11	2.0	2120	4	US-10-142-885-73	Sequence 73, Appl
706	11	2.0	1271	4	US-10-282-122A-63575	Sequence 63575, A	779	11	2.0	2120	4	US-10-158-790-73	Sequence 73, Appl
707	11	2.0	1313	4	US-10-408-765A-2180	Sequence 2180, Ap	780	11	2.0	2120	4	US-10-137-871-73	Sequence 73, Appl
708	11	2.0	1428	6	US-11-097-143-37923	Sequence 37923, A	781	11	2.0	2120	4	US-10-140-923-73	Sequence 73, Appl
709	11	2.0	1468	6	US-11-097-143-22779	Sequence 22779, A	782	11	2.0	2120	4	US-10-141-756-73	Sequence 73, Appl
710	11	2.0	1591	4	US-10-073-912-13	Sequence 13, Appl	783	11	2.0	2120	4	US-10-141-759-73	Sequence 73, Appl
711	11	2.0	1682	4	US-10-755-889-132	Sequence 132, App	784	11	2.0	2120	4	US-10-140-805-73	Sequence 73, Appl
712	11	2.0	1701	4	US-10-123-155-29	Sequence 29, Appl	785	11	2.0	2120	4	US-10-140-864-73	Sequence 73, Appl
713	11	2.0	1701	4	US-10-146-731-29	Sequence 29, Appl	786	11	2.0	2128	4	US-10-184-644-171	Sequence 171, App
714	11	2.0	1701	4	US-10-140-472-29	Sequence 29, Appl	787	11	2.0	2128	4	US-10-184-634-171	Sequence 171, App
715	11	2.0	1701	4	US-10-141-761-29	Sequence 29, Appl	788	11	2.0	2197	4	US-10-123-155-313	Sequence 313, App
716	11	2.0	1701	4	US-10-142-885-29	Sequence 29, Appl	789	11	2.0	2197	4	US-10-146-731-313	Sequence 313, App
717	11	2.0	1701	4	US-10-158-790-29	Sequence 29, Appl	790	11	2.0	2197	4	US-10-140-472-313	Sequence 313, App
718	11	2.0	1701	4	US-10-137-871-29	Sequence 29, Appl	791	11	2.0	2197	4	US-10-141-759-313	Sequence 313, App
719	11	2.0	1701	4	US-10-140-923-29	Sequence 29, Appl	792	11	2.0	2197	4	US-10-142-885-313	Sequence 313, App
720	11	2.0	1701	4	US-10-141-756-29	Sequence 29, Appl	793	11	2.0	2197	4	US-10-158-790-313	Sequence 313, App
721	11	2.0	1701	4	US-10-141-759-29	Sequence 29, Appl	794	11	2.0	2197	4	US-10-137-871-313	Sequence 313, App
722	11	2.0	1701	4	US-10-140-805-29	Sequence 29, Appl	795	11	2.0	2197	4	US-10-140-923-313	Sequence 313, App
723	11	2.0	1701	4	US-10-140-864-29	Sequence 29, Appl	796	11	2.0	2197	4	US-10-141-756-313	Sequence 313, App
724	11	2.0	1825	4	US-10-123-155-155	Sequence 155, App	797	11	2.0	2197	4	US-10-141-759-313	Sequence 313, App
725	11	2.0	1825	4	US-10-146-731-155	Sequence 155, App	798	11	2.0	2197	4	US-10-140-805-313	Sequence 313, App
726	11	2.0	1825	4	US-10-140-472-155	Sequence 155, App	799	11	2.0	2197	4	US-10-140-864-313	Sequence 313, App
727	11	2.0	1825	4	US-10-141-761-155	Sequence 155, App	800	11	2.0	2265	4	US-10-184-644-607	Sequence 607, App
728	11	2.0	1825	4	US-10-142-885-155	Sequence 155, App	801	11	2.0	2265	4	US-10-184-634-607	Sequence 607, App
729	11	2.0	1825	4	US-10-158-790-155	Sequence 155, App	802	11	2.0	2296	6	US-11-097-143-13942	Sequence 13942, Ap
730	11	2.0	1825	4	US-10-137-871-155	Sequence 155, App	803	11	2.0	2344	5	US-10-732-923-13978	Sequence 13978, A
731	11	2.0	1825	4	US-10-140-923-155	Sequence 155, App	804	11	2.0	2344	4	US-10-123-155-327	Sequence 327, App
732	11	2.0	1825	4	US-10-141-756-155	Sequence 155, App	805	11	2.0	2454	4	US-10-146-731-327	Sequence 327, App
733	11	2.0	1825	4	US-10-141-759-155	Sequence 155, App	806	11	2.0	2454	4	US-10-140-472-327	Sequence 327, App
734	11	2.0	1825	4	US-10-140-805-155	Sequence 155, App	807	11	2.0	2454	4	US-10-141-761-327	Sequence 327, App
735	11	2.0	1825	4	US-10-140-864-155	Sequence 155, App	808	11	2.0	2454	4	US-10-142-885-327	Sequence 327, App
736	11	2.0	1837	4	US-10-123-155-243	Sequence 243, App	809	11	2.0	2454	4	US-10-158-790-327	Sequence 327, App
737	11	2.0	1837	4	US-10-146-731-243	Sequence 243, App	810	11	2.0	2454	4	US-10-137-871-327	Sequence 327, App
738	11	2.0	1837	4	US-10-140-472-243	Sequence 243, App	811	11	2.0	2454	4	US-10-140-923-327	Sequence 327, App
739	11	2.0	1837	4	US-10-141-761-243	Sequence 243, App	812	11	2.0	2454	4	US-10-141-756-327	Sequence 327, App
740	11	2.0	1837	4	US-10-142-885-243	Sequence 243, App	813	11	2.0	2454	4	US-10-141-759-327	Sequence 327, App
741	11	2.0	1837	4	US-10-158-790-243	Sequence 243, App	814	11	2.0	2454	4	US-10-140-805-327	Sequence 327, App
742	11	2.0	1837	4	US-10-137-871-243	Sequence 243, App	815	11	2.0	2454	4	US-10-140-864-327	Sequence 327, App
743	11	2.0	1837	4	US-10-140-923-243	Sequence 243, App	816	11	2.0	2478	4	US-10-184-644-107	Sequence 107, App
744	11	2.0	1837	4	US-10-141-756-243	Sequence 243, App	817	11	2.0	2478	4	US-10-184-634-107	Sequence 107, App
745	11	2.0	1837	4	US-10-141-759-243	Sequence 243, App	818	11	2.0	2478	4	US-10-123-155-85	Sequence 85, Appl
746	11	2.0	1837	4	US-10-140-805-243	Sequence 243, App	819	11	2.0	2750	4	US-10-146-731-85	Sequence 85, Appl
747	11	2.0	1837	4	US-10-140-864-243	Sequence 243, App	820	11	2.0	2750	4	US-10-140-472-85	Sequence 85, Appl
748	11	2.0	1915	4	US-10-184-644-275	Sequence 275, App	821	11	2.0	2750	4	US-10-141-761-85	Sequence 85, Appl
749	11	2.0	1915	4	US-10-184-634-275	Sequence 275, App	822	11	2.0	2750	4	US-10-142-885-85	Sequence 85, Appl
750	11	2.0	2014	4	US-10-123-155-51	Sequence 51, Appl	823	11	2.0	2750	4	US-10-158-790-85	Sequence 85, Appl
751	11	2.0	2014	4	US-10-146-731-51	Sequence 51, Appl	824	11	2.0	2750	4	US-10-137-871-85	Sequence 85, Appl
752	11	2.0	2014	4	US-10-140-472-51	Sequence 51, Appl	825	11	2.0	2750	4	US-10-140-923-85	Sequence 85, Appl
753	11	2.0	2014	4	US-10-141-761-51	Sequence 51, Appl	826	11	2.0	2750	4	US-10-141-756-85	Sequence 85, Appl
754	11	2.0	2014	4	US-10-142-885-51	Sequence 51, Appl	827	11	2.0	2750	4	US-10-141-759-85	Sequence 85, Appl
755	11	2.0	2014	4	US-10-158-790-51	Sequence 51, Appl	828	11	2.0	2750	4	US-10-140-805-85	Sequence 85, Appl
756	11	2.0	2014	4	US-10-137-871-51	Sequence 51, Appl	829	11	2.0	2750	4	US-10-140-864-85	Sequence 85, Appl
757	11	2.0	2014	4	US-10-140-923-51	Sequence 51, Appl	830	11	2.0	2786	4	US-10-184-644-307	Sequence 307, App

831	11	2.0	2786	4	US-10-184-634-307	Sequence 307, App	904	11	2.0	4563	4	US-10-184-644-311	Sequence 311, App
832	11	2.0	2800	4	US-10-417-426-15	Sequence 15, App1	905	11	2.0	4563	4	US-10-184-634-311	Sequence 311, App
833	11	2.0	2837	4	US-10-123-155-229	Sequence 229, App	906	11	2.0	4640	4	US-10-184-634-75	Sequence 75, App1
834	11	2.0	2837	4	US-10-146-731-229	Sequence 229, App	907	11	2.0	4640	4	US-10-184-634-75	Sequence 75, App1
835	11	2.0	2837	4	US-10-140-472-229	Sequence 229, App	908	11	2.0	4640	4	US-10-184-634-75	Sequence 75, App1
836	11	2.0	2837	4	US-10-141-761-229	Sequence 229, App	909	11	2.0	35346	5	US-10-073-912-17	Sequence 17, App1
837	11	2.0	2837	4	US-10-142-885-229	Sequence 229, App	910	11	2.0	35823	5	US-10-874-049-1	Sequence 1, App1
838	11	2.0	2837	4	US-10-158-790-229	Sequence 229, App	911	11	1.8	10	3	US-09-573-822C-524	Sequence 524, App
839	11	2.0	2837	4	US-10-137-871-229	Sequence 229, App	912	10	1.8	10	3	US-09-573-822C-526	Sequence 526, App
840	11	2.0	2837	4	US-10-140-923-229	Sequence 229, App	913	10	1.8	10	3	US-09-573-822C-534	Sequence 534, App
841	11	2.0	2837	4	US-10-141-756-229	Sequence 229, App	914	10	1.8	10	3	US-09-573-822C-536	Sequence 536, App
842	11	2.0	2837	4	US-10-141-759-229	Sequence 229, App	915	10	1.8	10	3	US-09-573-822C-550	Sequence 550, App
843	11	2.0	2837	4	US-10-140-805-229	Sequence 229, App	916	10	1.8	10	3	US-09-573-822C-552	Sequence 552, App
844	11	2.0	2837	4	US-10-140-864-229	Sequence 229, App	917	10	1.8	10	3	US-09-573-822C-560	Sequence 560, App
845	11	2.0	2849	4	US-10-123-155-285	Sequence 285, App	918	10	1.8	10	3	US-09-573-822C-562	Sequence 562, App
846	11	2.0	2849	4	US-10-146-731-285	Sequence 285, App	919	10	1.8	10	3	US-09-573-822C-562	Sequence 562, App
847	11	2.0	2849	4	US-10-140-472-285	Sequence 285, App	920	10	1.8	11	6	US-11-103-356A-14	Sequence 14, App1
848	11	2.0	2849	4	US-10-141-761-285	Sequence 285, App	921	10	1.8	104	4	US-10-424-599-196147	Sequence 196147, App
849	11	2.0	2849	4	US-10-142-885-285	Sequence 285, App	922	10	1.8	103	4	US-10-424-599-195398	Sequence 195398, App
850	11	2.0	2849	4	US-10-158-790-285	Sequence 285, App	923	10	1.8	115	5	US-10-450-763-56856	Sequence 56856, App
851	11	2.0	2849	4	US-10-137-871-285	Sequence 285, App	924	10	1.8	125	4	US-10-425-115-202978	Sequence 202978, App
852	11	2.0	2849	4	US-10-140-923-285	Sequence 285, App	925	10	1.8	133	6	US-11-172-740-1477	Sequence 1477, App
853	11	2.0	2849	4	US-10-141-756-285	Sequence 285, App	926	10	1.8	151	4	US-10-424-599-192174	Sequence 192174, App
854	11	2.0	2849	4	US-10-141-759-285	Sequence 285, App	927	10	1.8	158	4	US-10-425-114-50482	Sequence 50482, App
855	11	2.0	2849	4	US-10-140-805-285	Sequence 285, App	928	10	1.8	171	4	US-10-437-963-203069	Sequence 203069, App
856	11	2.0	2849	4	US-10-140-864-285	Sequence 285, App	929	10	1.8	195	6	US-11-097-143-34818	Sequence 34818, App
857	11	2.0	3003	4	US-10-184-644-51	Sequence 51, App1	930	10	1.8	213	4	US-10-424-599-240414	Sequence 240414, App
858	11	2.0	3038	4	US-10-184-634-51	Sequence 51, App1	931	10	1.8	222	6	US-10-424-599-283027	Sequence 283027, App
859	11	2.0	3038	4	US-10-184-644-261	Sequence 261, App	932	10	1.8	222	6	US-11-096-568A-8037	Sequence 8037, App
860	11	2.0	3313	4	US-10-184-634-261	Sequence 261, App	933	10	1.8	224	5	US-10-732-923-5552	Sequence 5552, App
861	11	2.0	3313	4	US-10-184-634-77	Sequence 77, App1	934	10	1.8	244	4	US-10-424-599-176078	Sequence 176078, App
862	11	2.0	3313	4	US-10-123-155-35	Sequence 35, App1	935	10	1.8	253	4	US-10-278-173-44	Sequence 17344, App
863	11	2.0	3371	4	US-10-146-731-35	Sequence 35, App1	936	10	1.8	253	4	US-10-412-699B-18	Sequence 18, App1
864	11	2.0	3371	4	US-10-140-923-35	Sequence 35, App1	937	10	1.8	253	5	US-11-097-143-36735	Sequence 36735, App
865	11	2.0	3371	4	US-10-142-885-35	Sequence 35, App1	938	10	1.8	254	6	US-10-425-115-356052	Sequence 356052, App
866	11	2.0	3371	4	US-10-141-761-35	Sequence 35, App1	939	10	1.8	255	5	US-10-425-114-38198	Sequence 38198, App
867	11	2.0	3371	4	US-10-158-790-35	Sequence 35, App1	940	10	1.8	259	6	US-11-172-740-947	Sequence 947, App
868	11	2.0	3371	4	US-10-137-871-35	Sequence 35, App1	941	10	1.8	265	4	US-10-437-963-125866	Sequence 125866, App
869	11	2.0	3371	4	US-10-140-923-35	Sequence 35, App1	942	10	1.8	307	6	US-10-437-963-197107	Sequence 197107, App
870	11	2.0	3371	4	US-10-141-759-35	Sequence 35, App1	943	10	1.8	316	6	US-11-087-099-7594	Sequence 7594, App
871	11	2.0	3371	4	US-10-141-756-35	Sequence 35, App1	944	10	1.8	335	4	US-10-295-403-58	Sequence 58, App1
872	11	2.0	3371	4	US-10-140-805-35	Sequence 35, App1	945	10	1.8	335	4	US-10-225-066A-222	Sequence 222, App
873	11	2.0	3371	4	US-10-140-864-35	Sequence 35, App1	946	10	1.8	335	4	US-10-374-780A-2066	Sequence 2066, App
874	11	2.0	3401	4	US-10-184-644-411	Sequence 411, App	947	10	1.8	335	4	US-10-412-699B-628	Sequence 628, App
875	11	2.0	3401	4	US-10-184-644-411	Sequence 411, App	948	10	1.8	335	4	US-10-412-699B-1820	Sequence 1820, App
876	11	2.0	3554	4	US-10-123-155-537	Sequence 537, App	949	10	1.8	335	5	US-10-225-066A-222	Sequence 222, App
877	11	2.0	3554	4	US-10-146-731-537	Sequence 537, App	950	10	1.8	335	5	US-10-424-599-165226	Sequence 165226, App
878	11	2.0	3554	4	US-10-140-472-537	Sequence 537, App	951	10	1.8	471	4	US-10-437-963-125866	Sequence 125866, App
879	11	2.0	3554	4	US-10-141-759-537	Sequence 537, App	952	10	1.8	489	4	US-10-437-963-197107	Sequence 197107, App
880	11	2.0	3554	4	US-10-141-761-537	Sequence 537, App	953	10	1.8	491	4	US-10-437-963-197107	Sequence 197107, App
881	11	2.0	3554	4	US-10-142-885-537	Sequence 537, App	954	10	1.8	525	4	US-10-719-385-25	Sequence 25, App1
882	11	2.0	3554	4	US-10-158-790-537	Sequence 537, App	955	10	1.8	525	4	US-10-032-58B-7161	Sequence 7161, App
883	11	2.0	3554	4	US-10-137-871-537	Sequence 537, App	956	10	1.8	528	4	US-10-282-122A-72941	Sequence 72941, App
884	11	2.0	3554	4	US-10-140-923-537	Sequence 537, App	957	10	1.8	533	3	US-09-804-551B-28	Sequence 28, App1
885	11	2.0	3554	4	US-10-141-756-537	Sequence 537, App	958	10	1.8	535	4	US-10-123-155-525	Sequence 525, App
886	11	2.0	3554	4	US-10-140-805-537	Sequence 537, App	959	10	1.8	535	4	US-10-146-731-525	Sequence 525, App
887	11	2.0	3554	4	US-10-141-759-537	Sequence 537, App	960	10	1.8	535	4	US-10-140-472-535	Sequence 525, App
888	11	2.0	3554	4	US-10-140-864-537	Sequence 537, App	961	10	1.8	535	4	US-10-141-756-525	Sequence 525, App
889	11	2.0	4038	4	US-10-417-426-14	Sequence 14, App1	962	10	1.8	535	4	US-10-142-885-525	Sequence 525, App
890	11	2.0	4060	4	US-10-123-155-197	Sequence 197, App	963	10	1.8	535	4	US-10-158-792-525	Sequence 525, App
891	11	2.0	4060	4	US-10-146-731-197	Sequence 197, App	964	10	1.8	535	4	US-10-137-871-525	Sequence 525, App
892	11	2.0	4060	4	US-10-140-472-197	Sequence 197, App	965	10	1.8	535	4	US-10-140-923-525	Sequence 525, App
893	11	2.0	4060	4	US-10-141-761-197	Sequence 197, App	966	10	1.8	535	4	US-10-141-756-525	Sequence 525, App
894	11	2.0	4060	4	US-10-142-885-197	Sequence 197, App	967	10	1.8	535	4	US-10-141-759-525	Sequence 525, App
895	11	2.0	4060	4	US-10-158-790-197	Sequence 197, App	968	10	1.8	535	4	US-10-140-805-525	Sequence 525, App
896	11	2.0	4060	4	US-10-137-871-197	Sequence 197, App	969	10	1.8	535	4	US-10-140-864-525	Sequence 525, App
897	11	2.0	4060	4	US-10-140-923-197	Sequence 197, App	970	10	1.8	579	6	US-11-097-143-44485	Sequence 44485, App
898	11	2.0	4060	4	US-10-141-756-197	Sequence 197, App	971	10	1.8	625	4	US-10-425-114-55660	Sequence 55660, App
899	11	2.0	4060	4	US-10-141-759-197	Sequence 197, App	972	10	1.8	642	4	US-10-424-599-162645	Sequence 162645, App
900	11	2.0	4060	4	US-10-140-805-197	Sequence 197, App	973	10	1.8	648	5	US-10-732-923-10742	Sequence 10742, App
901	11	2.0	4060	4	US-10-140-864-197	Sequence 197, App	974	10	1.8	659	4	US-10-437-963-160045	Sequence 160045, App
902	11	2.0	4407	4	US-10-184-644-351	Sequence 351, App	975	10	1.8	788	5	US-10-489-423-40	Sequence 40, App1
903	11	2.0	4407	4	US-10-184-634-351	Sequence 351, App	976	10	1.8	788	6	US-11-097-143-8541	Sequence 8541, App


```
977 10 1.8 842 4 US-10-032-585-7311 Sequence 7311, App
978 10 1.8 904 4 US-10-437-963-136970 Sequence 136970,
979 10 1.8 948 4 US-10-268-473-4 Sequence 177, App
980 10 1.8 959 4 US-10-123-155-177 Sequence 177, App
981 10 1.8 959 4 US-10-146-731-177 Sequence 177, App
982 10 1.8 959 4 US-10-140-472-177 Sequence 177, App
983 10 1.8 959 4 US-10-141-761-177 Sequence 177, App
984 10 1.8 959 4 US-10-142-885-177 Sequence 177, App
985 10 1.8 959 4 US-10-158-792-177 Sequence 177, App
986 10 1.8 959 4 US-10-137-871-177 Sequence 177, App
987 10 1.8 959 4 US-10-140-923-177 Sequence 177, App
988 10 1.8 959 4 US-10-141-756-177 Sequence 177, App
989 10 1.8 959 4 US-10-141-756-177 Sequence 177, App
990 10 1.8 959 4 US-10-140-805-177 Sequence 177, App
991 10 1.8 959 4 US-10-140-864-177 Sequence 10, App
992 10 1.8 1009 4 US-10-654-416-10 Sequence 12, App
993 10 1.8 1017 4 US-10-654-416-12 Sequence 14, App
994 10 1.8 1023 4 US-10-654-416-14 Sequence 6, App
995 10 1.8 1027 4 US-10-654-416-8 Sequence 8, App
996 10 1.8 1029 4 US-10-654-416-6 Sequence 23, App
997 10 1.8 1111 4 US-10-719-385-23 Sequence 22513, A
998 10 1.8 1208 5 US-10-732-923-22513 Sequence 509, App
999 10 1.8 1281 4 US-10-123-155-509 Sequence 509, App
1000 10 1.8 1281 4 US-10-146-731-509 Sequence 509, App
```

ALIGNMENTS

```
RESULT 1
US-11-134-563-12
; Sequence 12, Application US/11134563
; Publication No. US20050287569A1
GENERAL INFORMATION:
; APPLICANT: Leong, John M.
; APPLICANT: Campbellone, Kenneth G.
; TITLE OF INVENTION: ESPFNUCLEIC ACIDS AND PROTEINS AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: 07917-280001
; CURRENT APPLICATION NUMBER: US/11/134,563
; CURRENT FILING DATE: 2005-05-20
; PRIOR APPLICATION NUMBER: US 60/573,600
; PRIOR FILING DATE: 2004-05-20
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 558
; TYPE: PRT
; ORGANISM: Escherichia coli
US-11-134-563-12
```

```
Query Match 100.0%; Score 558; DB 6; Length 558;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 558; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MPICGLGNPNVNSIPAPPLPSQTDGAGRGOLINSTGPIGSRALFTPVNSMADSGD 60
DB 1 MPICGLGNPNVNSIPAPPLPSQTDGAGRGOLINSTGPIGSRALFTPVNSMADSGD 60
QY 61 NRASDVPGIPVNPMLAASEITLNDGFEVLHDHGPLDLTLNRQIGSSVRFVETOEDGKHIA 120
DB 61 NRASDVPGIPVNPMLAASEITLNDGFEVLHDHGPLDLTLNRQIGSSVRFVETOEDGKHIA 120
QY 121 VQGRNGVETSVVLSQOEVARLOSIDPEKDKFVFTGGGAGGHAMVTASDITTEARORIL 180
DB 121 VQGRNGVETSVVLSQOEVARLOSIDPEKDKFVFTGGGAGGHAMVTASDITTEARORIL 180
QY 181 ELLEPKGTGESKGAGSGKGVGLRBSNSGAENTTETOTSTSSIRSDPKMLALGYAT 240
DB 181 ELLEPKGTGESKGAGSGKGVGLRBSNSGAENTTETOTSTSSIRSDPKMLALGYAT 240
QY 241 GLIGLAATGIVQALALTPEDPSPTTDDPAASATETATRDQLTEAFQNPNDQKVNIDE 300
DB 241 GLIGLAATGIVQALALTPEDPSPTTDDPAASATETATRDQLTEAFQNPNDQKVNIDE 300
```

```
DB 241 GLIGLAATGIVQALALTPEDPSPTTDDPAASATETATRDQLTEAFQNPNDQKVNIDE 300
QY 301 LGNAIPSGVLKDDPVANIEBOAKAAGEBAKQOAIENNAQOKKYBOQAKROELKVSSG 360
DB 301 LGNAIPSGVLKDDPVANIEBOAKAAGEBAKQOAIENNAQOKKYBOQAKROELKVSSG 360
QY 361 AGYGLSGALILGGGIGVAVTAALHRKNQPVBOTTTTTTTTTSARIVENKPAANTPAQG 420
DB 361 AGYGLSGALILGGGIGVAVTAALHRKNQPVBOTTTTTTTTTSARIVENKPAANTPAQG 420
QY 421 NVDTFGSEDTHMESRRSSMASTSTFFDTSIGTVONPADVKTSLHDSQVPTSNSTSVQ 480
DB 421 NVDTFGSEDTHMESRRSSMASTSTFFDTSIGTVONPADVKTSLHDSQVPTSNSTSVQ 480
QY 481 NMGNDSVYVSTIOHPPTDITDNGARLGNPSAGIQTSTYARLASGGRHDMGGLTGGSN 540
DB 481 NMGNDSVYVSTIOHPPTDITDNGARLGNPSAGIQTSTYARLASGGRHDMGGLTGGSN 540
QY 541 SAVNTSNPNPAPGSHRFV 558
DB 541 SAVNTSNPNPAPGSHRFV 558
```

```
RESULT 2
US-11-052-554A-20
; Sequence 20, Application US/11052554A
; Publication No. US20050288866A1
GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; TITLE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 20
; LENGTH: 558
; TYPE: PRT
; ORGANISM: Escherichia coli 0157:H7
US-11-052-554A-20
```

```
Query Match 100.0%; Score 558; DB 6; Length 558;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 558; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MPICGLGNPNVNSIPAPPLPSQTDGAGRGOLINSTGPIGSRALFTPVNSMADSGD 60
DB 1 MPICGLGNPNVNSIPAPPLPSQTDGAGRGOLINSTGPIGSRALFTPVNSMADSGD 60
QY 61 NRASDVPGIPVNPMLAASEITLNDGFEVLHDHGPLDLTLNRQIGSSVRFVETOEDGKHIA 120
DB 61 NRASDVPGIPVNPMLAASEITLNDGFEVLHDHGPLDLTLNRQIGSSVRFVETOEDGKHIA 120
QY 121 VQGRNGVETSVVLSQOEVARLOSIDPEKDKFVFTGGGAGGHAMVTASDITTEARORIL 180
DB 121 VQGRNGVETSVVLSQOEVARLOSIDPEKDKFVFTGGGAGGHAMVTASDITTEARORIL 180
QY 181 ELLEPKGTGESKGAGSGKGVGLRBSNSGAENTTETOTSTSSIRSDPKMLALGYAT 240
DB 181 ELLEPKGTGESKGAGSGKGVGLRBSNSGAENTTETOTSTSSIRSDPKMLALGYAT 240
QY 241 GLIGLAATGIVQALALTPEDPSPTTDDPAASATETATRDQLTEAFQNPNDQKVNIDE 300
DB 241 GLIGLAATGIVQALALTPEDPSPTTDDPAASATETATRDQLTEAFQNPNDQKVNIDE 300
QY 301 LGNAIPSGVLKDDPVANIEBOAKAAGEBAKQOAIENNAQOKKYBOQAKROELKVSSG 360
DB 301 LGNAIPSGVLKDDPVANIEBOAKAAGEBAKQOAIENNAQOKKYBOQAKROELKVSSG 360
```



```
Qy 361 AGYGLSGALLGGIGVAVTAALHRKNQPVETTTTTTTTTTTSARTVENKPNANTPAOG 420
|||
Db 361 AGYGLSGALLGGIGVAVTAALHRKNQPVETTTTTTTTTTTSARTVENKPNANTPAOG 420

Qy 421 NVDTPOSEDMESRRSSMASTSTFPDTSIGTVQNPYADVKTSLHDSQVPTSNSTSVQ 480
|||
Db 421 NVDTPOSEDMESRRSSMASTSTFPDTSIGTVQNPYADVKTSLHDSQVPTSNSTSVQ 480

Qy 481 NMGNTSVYVSTTQHPPTDTTNGARLLGNPSAGIOSTVARLALSGGLRHDMGGLTGGSN 540
|||
Db 481 NMGNTSVYVSTTQHPPTDTTNGARLLGNPSAGIOSTVARLALSGGLRHDMGGLTGGSN 540

Qy 541 SAVNTSNPPAPGSHRFV 558
|||
Db 541 SAVNTSNPPAPGSHRFV 558

RESULT 3
US-10-470-048B-73
; Sequence 73, Application US/10470048B
; Publication No. US2005003744A1
; GENERAL INFORMATION:
; APPLICANT: MEINKE ET AL.
; TITLE OF INVENTION: A METHOD FOR IDENTIFICATION, ISOLATION AND PRODUCTION OF
; FILE REFERENCE: SONN:035US
; CURRENT APPLICATION NUMBER: US/10/470,048B
; CURRENT FILING DATE: 2003-07-25
; NUMBER OF SEQ ID NOS: 603
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 73
; LENGTH: 166
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-470-048B-73

Query Match 2.3%; Score 13; DB 5; Length 166;
Best Local Similarity 100.0%; Pred. No. 0.0031;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 392 QTTTTTTTTTTS 404
|||
Db 43 QTTTTTTTTTTS 55

RESULT 4
US-10-424-599-207017
; Sequence 207017, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 207017
; LENGTH: 107
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(107)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_28963C.1.pep
US-10-424-599-207017
```

```
Query Match 2.2%; Score 12; DB 4; Length 107;
Best Local Similarity 100.0%; Pred. No. 0.016;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 392 QTTTTTTTTTTS 403
|||
Db 52 QTTTTTTTTTTS 63

RESULT 5
US-10-425-115-366802
; Sequence 366802, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 366802
; LENGTH: 278
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_9768C.1.pep
US-10-425-115-366802

Query Match 2.2%; Score 12; DB 4; Length 278;
Best Local Similarity 100.0%; Pred. No. 0.042;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 393 TTTTTTTTTTTS 404
|||
Db 21 TTTTTTTTTTTS 32

RESULT 6
US-10-017-161-1796
; Sequence 1796, Application US/10017161
; Publication No. US20030143668A1
; GENERAL INFORMATION:
; APPLICANT: SUMA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABURATANI, HIROYUKI
; TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 08435/0152
; CURRENT APPLICATION NUMBER: US/10/017,161
; CURRENT FILING DATE: 2002-12-18
; PRIOR APPLICATION NUMBER: JP 2001/246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2430
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1796
; LENGTH: 361
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-017-161-1796

Query Match 2.2%; Score 12; DB 4; Length 361;
Best Local Similarity 100.0%; Pred. No. 0.053;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 393 TTTTTTTTTTTS 404
|||
Db 334 TTTTTTTTTTTS 345
```

```
RESULT 7
US-10-292-798-1452
; Sequence 1452, Application US/10292798
; Publication No. US20030235833A1
; GENERAL INFORMATION:
; APPLICANT: SUMA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABURATANI, HIROYUKI
; TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS
; FILE REFERENCE: 08435/166
; CURRENT APPLICATION NUMBER: US/10/292,798
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 10/017,161
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: JP 2001-246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2070
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1452
; LENGTH: 361
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-292-798-1452

Query Match
Best Local Similarity 100.0%; Score 12; DB 4; Length 361;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 393 TTTTTTTTTTS 404
Db 334 TTTTTTTTTTS 345

RESULT 8
US-10-424-599-144079
; Sequence 144079, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5322)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 144079
; LENGTH: 367
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_101115C.1.pep
US-10-424-599-144079

Query Match
Best Local Similarity 100.0%; Score 12; DB 4; Length 367;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 393 TTTTTTTTTTS 404
Db 241 TTTTTTTTTTS 252

RESULT 9
US-11-097-143-30426
; Sequence 30426, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.

RESULT 10
US-10-741-849-7161
; Sequence 7161, Application US/10741849
; Publication No. US20050019931A1
; GENERAL INFORMATION:
; APPLICANT: Roemer, Terry
; APPLICANT: Jiang, Bo
; APPLICANT: Boone, Charles
; APPLICANT: Bussey, Howard
; TITLE OF INVENTION: Nucleic Acids Encoding Anti-fungal Drug Targets and Methods of
; TITLE OF INVENTION: Use
; FILE REFERENCE: 10182-023-999
; CURRENT APPLICATION NUMBER: US/10/741,849
; CURRENT FILING DATE: 2003-12-19
; PRIOR APPLICATION NUMBER: US 60/434,832
; PRIOR FILING DATE: 2002-12-19
; NUMBER OF SEQ ID NOS: 8000
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7161
; LENGTH: 785
; TYPE: PRT
; ORGANISM: Candida albicans
US-10-741-849-7161

Query Match
Best Local Similarity 100.0%; Score 12; DB 5; Length 785;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 393 TTTTTTTTTTS 404
Db 16 TTTTTTTTTTS 27

RESULT 11
US-11-097-143-30426
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; TITLE OF INVENTION: DROSOPHILA GENES.
; FILE REFERENCE: C1000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30426
; LENGTH: 561
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-11-097-143-30426

Query Match
Best Local Similarity 100.0%; Score 12; DB 6; Length 561;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 393 TTTTTTTTTTS 404
Db 112 TTTTTTTTTTS 123
```

```
RESULT 10
US-10-741-849-7161
; Sequence 7161, Application US/10741849
; Publication No. US20050019931A1
; GENERAL INFORMATION:
; APPLICANT: Roemer, Terry
; APPLICANT: Jiang, Bo
; APPLICANT: Boone, Charles
; APPLICANT: Bussey, Howard
; TITLE OF INVENTION: Nucleic Acids Encoding Anti-fungal Drug Targets and Methods of
; TITLE OF INVENTION: Use
; FILE REFERENCE: 10182-023-999
; CURRENT APPLICATION NUMBER: US/10/741,849
; CURRENT FILING DATE: 2003-12-19
; PRIOR APPLICATION NUMBER: US 60/434,832
; PRIOR FILING DATE: 2002-12-19
; NUMBER OF SEQ ID NOS: 8000
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7161
; LENGTH: 785
; TYPE: PRT
; ORGANISM: Candida albicans
US-10-741-849-7161

Query Match
Best Local Similarity 100.0%; Score 12; DB 5; Length 785;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 393 TTTTTTTTTTS 404
Db 16 TTTTTTTTTTS 27

RESULT 11
US-11-097-143-30426
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; TITLE OF INVENTION: DROSOPHILA GENES.
; FILE REFERENCE: C1000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30426
; LENGTH: 561
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-11-097-143-30426

Query Match
Best Local Similarity 100.0%; Score 12; DB 6; Length 561;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 393 TTTTTTTTTTS 404
Db 112 TTTTTTTTTTS 123
```

```
US-11-188-298-5674
; Sequence 5674, Application US/11188298
; Publication No. US20060075522A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: GENES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53452)B
; CURRENT APPLICATION NUMBER: US/11/188,298
; PRIOR FILING DATE: 2005-07-22
; PRIOR APPLICATION NUMBER: 60/592,978
; PRIOR FILING DATE: 2004-07-31
; NUMBER OF SEQ ID NOS: 22569
; SEQ ID NO 5674
; LENGTH: 802
; TYPE: PRT
; ORGANISM: Ruminococcus flavefaciens
US-11-188-298-5674

Query Match      2.2%; Score 12; DB 6; Length 802;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      393 TTTT TTTT TTTT TTTT 404
      |||||
Db      533 TTTT TTTT TTTT TTTT 544

RESULT 12
US-11-188-298-20820
; Sequence 20820, Application US/11188298
; Publication No. US20060075522A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: GENES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53452)B
; CURRENT APPLICATION NUMBER: US/11/188,298
; PRIOR FILING DATE: 2005-07-22
; PRIOR APPLICATION NUMBER: 60/592,978
; PRIOR FILING DATE: 2004-07-31
; NUMBER OF SEQ ID NOS: 22569
; SEQ ID NO 20820
; LENGTH: 802
; TYPE: PRT
; ORGANISM: Ruminococcus flavefaciens
US-11-188-298-20820

Query Match      2.2%; Score 12; DB 6; Length 802;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      393 TTTT TTTT TTTT TTTT 404
      |||||
Db      533 TTTT TTTT TTTT TTTT 544

RESULT 13
US-10-363-946-2
; Sequence 2, Application US/10363946
; Publication No. US20050037436A1
; GENERAL INFORMATION:
; APPLICANT: Bayer Aktiengesellschaft
; APPLICANT: von Samson-Himmelsfjerna, Georg
; APPLICANT: Harder, Achim
; APPLICANT: Wunderlich, Frank
; APPLICANT: Schmitt-Wrede, Hans-Peter
; APPLICANT: Saeger, Beate
; TITLE OF INVENTION: Test Systems and the Use Thereof for Identifying and
; FILE REFERENCE: Lea 34 240
; CURRENT APPLICATION NUMBER: US/10/363,946
; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: PCT/EP01/09771
; PRIOR FILING DATE: 2001-08-24
```

```
; PRIOR APPLICATION NUMBER: DE100 44 098.3
; PRIOR FILING DATE: 2000-09-07
; PRIOR APPLICATION NUMBER: DE100 53 785.5
; PRIOR FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 986
; TYPE: PRT
; ORGANISM: Haemophilus contortus
US-10-363-946-2

Query Match      2.2%; Score 12; DB 5; Length 986;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      393 TTTT TTTT TTTT TTTT 404
      |||||
Db      133 TTTT TTTT TTTT TTTT 144

RESULT 14
US-11-097-143-19809
; Sequence 19809, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; TITLE OF INVENTION: DROSOPHILA GENES.
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; PRIOR FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19809
; LENGTH: 1026
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-11-097-143-19809

Query Match      2.2%; Score 12; DB 6; Length 1026;
Best Local Similarity 100.0%; Pred. No. 0.14;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      393 TTTT TTTT TTTT TTTT 404
      |||||
Db      605 TTTT TTTT TTTT TTTT 616

RESULT 15
US-11-097-143-38604
; Sequence 38604, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
```

```
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; TITLE OF INVENTION: DROSOPHILA GENES.
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38604
; LENGTH: 1026
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-11-097-143-38604
```

```
Query Match 2.2%: Score 12; DB 6; Length 1026;
Best Local Similarity 100.0%; Pred. NO. 0.14;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 392 QTTTTTTTTT 403
| | | | | | | | | |
Db 806 QTTTTTTTTT 817
```

Search completed: August 1, 2006, 22:53:43
Job time : 195 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: August 1, 2006, 22:50:36 ; Search time 32 Seconds
(without alignments)
1151.645 Million cell updates/sec

Title: US-09-189-415D-11

Perfect score: 558

Sequence: 1 MPIGNLGNPNVNNISIPAP.....SNSAVNTSNNPAPGSHRFV 558

Scoring table: OLIGO

Searched: 232337 seqs, 66044171 residues

Word size : 8

Total number of hits satisfying chosen parameters: 635

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database : Published Applications AA New.*

1: /EMC_Celerra_SIDS3/Ptodata/2/pubppaa/US09_NEW_PUB.rep.*
2: /EMC_Celerra_SIDS3/Ptodata/2/pubppaa/US06_NEW_PUB.rep.*
3: /EMC_Celerra_SIDS3/Ptodata/2/pubppaa/US07_NEW_PUB.rep.*
4: /EMC_Celerra_SIDS3/Ptodata/2/pubppaa/US08_NEW_PUB.rep.*
5: /EMC_Celerra_SIDS3/Ptodata/2/pubppaa/PCT_NEW_PUB.rep.*
6: /EMC_Celerra_SIDS3/Ptodata/2/pubppaa/US10_NEW_PUB.rep.*
7: /EMC_Celerra_SIDS3/Ptodata/2/pubppaa/US11_NEW_PUB.rep.*
8: /EMC_Celerra_SIDS3/Ptodata/2/pubppaa/US60_NEW_PUB.rep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	13	2.3	166	6	US-10-471-571A-782
2	12	2.2	264	6	US-10-953-349-34264
3	12	2.2	659	6	US-10-953-349-11680
4	12	2.2	1675	7	US-11-063-439-66
5	12	2.2	2823	7	US-11-063-439-24
6	12	2.2	3342	7	US-11-063-439-273
7	12	2.2	3445	7	US-11-063-439-68
8	12	2.2	3476	7	US-11-063-439-248
9	12	2.2	3481	7	US-11-063-439-86
10	12	2.2	3485	7	US-11-063-439-10
11	12	2.2	3485	7	US-11-063-439-59
12	12	2.2	3485	7	US-11-063-439-76
13	12	2.2	3486	7	US-11-063-439-45
14	12	2.2	3487	7	US-11-063-439-56
15	12	2.2	3488	7	US-11-063-439-42
16	12	2.2	3489	7	US-11-063-439-15
17	12	2.2	3498	7	US-11-063-439-2
18	12	2.2	3498	7	US-11-063-439-234
19	12	2.2	3501	7	US-11-063-439-75
20	12	2.2	3502	7	US-11-063-439-275
21	12	2.2	3503	7	US-11-063-439-23
22	12	2.2	3503	7	US-11-063-439-30
23	12	2.2	3504	7	US-11-063-439-247
24	12	2.2	3505	7	US-11-063-439-183
25	12	2.2	3505	7	US-11-063-439-208

26	12	2.2	3519	7	US-11-063-439-189	Sequence 189, App
27	12	2.2	3536	7	US-11-063-439-175	Sequence 175, App
28	12	2.2	3537	7	US-11-063-439-177	Sequence 177, App
29	12	2.2	3544	7	US-11-063-439-19	Sequence 19, App
30	12	2.2	3549	7	US-11-063-439-149	Sequence 149, App
31	12	2.2	3551	7	US-11-063-439-40	Sequence 40, App
32	12	2.2	3579	7	US-11-063-439-259	Sequence 259, App
33	12	2.2	3593	7	US-11-063-439-20	Sequence 20, App
34	12	2.2	3623	7	US-11-063-439-282	Sequence 282, App
35	12	2.2	3637	7	US-11-063-439-288	Sequence 288, App
36	12	2.2	3710	7	US-11-063-439-270	Sequence 270, App
37	12	2.2	3712	7	US-11-063-439-272	Sequence 272, App
38	12	2.2	3719	7	US-11-063-439-263	Sequence 263, App
39	11	2.0	181	7	US-11-056-3558-86782	Sequence 86782, A
40	11	2.0	181	7	US-11-056-3558-96259	Sequence 96259, A
41	11	2.0	186	7	US-11-056-3558-26451	Sequence 26451, A
42	11	2.0	203	7	US-11-056-3558-86781	Sequence 86781, A
43	11	2.0	208	7	US-11-056-3558-26450	Sequence 26450, A
44	11	2.0	229	7	US-11-246-999-46	Sequence 46, App
45	11	2.0	288	7	US-11-321-421-341	Sequence 341, App
46	11	2.0	288	7	US-11-321-421-344	Sequence 344, App
47	11	2.0	364	6	US-10-449-902-39786	Sequence 39786, A
48	11	2.0	364	7	US-11-246-999-39	Sequence 39, App
49	11	2.0	440	6	US-10-196-749-34	Sequence 34, App
50	11	2.0	440	7	US-11-296-092-61	Sequence 61, App
51	11	2.0	440	7	US-11-296-155-61	Sequence 61, App
52	11	2.0	442	7	US-11-246-999-136	Sequence 136, App
53	11	2.0	503	7	US-11-056-3558-38144	Sequence 38144, A
54	11	2.0	512	7	US-11-056-3558-38143	Sequence 38143, A
55	11	2.0	517	7	US-11-330-403-10087	Sequence 10087, A
56	11	2.0	518	7	US-11-330-403-7687	Sequence 7687, App
57	11	2.0	644	7	US-11-174-307B-4416	Sequence 4416, App
58	11	2.0	651	7	US-11-174-307B-3918	Sequence 3918, App
59	11	2.0	803	7	US-11-174-307B-2128	Sequence 2128, App
60	11	2.0	885	7	US-11-063-439-91	Sequence 91, App
61	11	2.0	997	7	US-11-174-307B-4404	Sequence 4404, App
62	11	2.0	1042	7	US-11-174-307B-2500	Sequence 2500, App
63	11	2.0	1143	7	US-11-174-307B-442	Sequence 492, App
64	11	2.0	1147	7	US-11-174-307B-2286	Sequence 2286, App
65	11	2.0	1181	7	US-11-174-307B-1288	Sequence 1288, App
66	11	2.0	1197	7	US-11-174-307B-1382	Sequence 1382, App
67	11	2.0	1206	7	US-11-174-307B-916	Sequence 916, App
68	11	2.0	1211	7	US-11-174-307B-2202	Sequence 2202, App
69	11	2.0	1211	7	US-11-174-307B-2480	Sequence 2480, App
70	11	2.0	1251	7	US-11-174-307B-4014	Sequence 4014, App
71	11	2.0	1254	7	US-11-174-307B-6686	Sequence 6686, App
72	11	2.0	1261	7	US-11-174-307B-668	Sequence 668, App
73	11	2.0	1262	7	US-11-063-439-82	Sequence 82, App
74	11	2.0	1296	7	US-11-174-307B-2080	Sequence 2080, App
75	11	2.0	1337	7	US-11-174-307B-376	Sequence 376, App
76	11	2.0	1351	7	US-11-174-307B-2808	Sequence 2808, App
77	11	2.0	1358	7	US-11-174-307B-2640	Sequence 2640, App
78	11	2.0	1375	7	US-11-174-307B-2276	Sequence 2276, App
79	11	2.0	1408	7	US-11-174-307B-2196	Sequence 2196, App
80	11	2.0	1424	7	US-11-174-307B-2542	Sequence 2542, App
81	11	2.0	1439	7	US-11-174-307B-3364	Sequence 3364, App
82	11	2.0	1461	7	US-11-174-307B-2824	Sequence 2824, App
83	11	2.0	1469	7	US-11-174-307B-1456	Sequence 1456, App
84	11	2.0	1471	7	US-11-174-307B-1160	Sequence 1160, App
85	11	2.0	1487	7	US-11-174-307B-696	Sequence 696, App
86	11	2.0	1502	7	US-11-174-307B-1490	Sequence 1490, App
87	11	2.0	1521	7	US-11-174-307B-1002	Sequence 1002, App
88	11	2.0	1523	7	US-11-174-307B-2716	Sequence 2716, App
89	11	2.0	1601	7	US-11-174-307B-2364	Sequence 2364, App
90	11	2.0	1692	7	US-11-174-307B-2858	Sequence 2858, App
91	11	2.0	1708	7	US-11-174-307B-2136	Sequence 2136, App
92	11	2.0	1894	7	US-11-063-439-88	Sequence 88, App
93	11	2.0	2068	7	US-11-174-307B-1032	Sequence 1032, App
94	11	2.0	2072	7	US-11-063-439-278	Sequence 278, App
95	11	2.0	2236	7	US-11-063-439-249	Sequence 249, App
96	11	2.0	2394	7	US-11-174-307B-2722	Sequence 2722, App
97	11	2.0	2504	7	US-11-063-439-115	Sequence 115, App
98	11	2.0	2523	7	US-11-063-439-51	Sequence 51, App

99	11	2.0	2537	7	US-11-063-439-43	Sequence 43, Appl	172	11	2.0	3502	7	US-11-063-439-144	Sequence 144, App
100	11	2.0	2539	7	US-11-063-439-62	Sequence 62, Appl	173	11	2.0	3502	7	US-11-063-439-160	Sequence 160, App
101	11	2.0	2576	7	US-11-063-439-223	Sequence 223, App	174	11	2.0	3502	7	US-11-063-439-184	Sequence 184, App
102	11	2.0	2662	7	US-11-174-307B-2676	Sequence 2676, App	175	11	2.0	3502	7	US-11-063-439-186	Sequence 186, App
103	11	2.0	3031	7	US-11-063-439-108	Sequence 108, Appl	176	11	2.0	3502	7	US-11-063-439-236	Sequence 236, App
104	11	2.0	3230	7	US-11-063-439-3	Sequence 3, Appl1	177	11	2.0	3503	7	US-11-063-439-140	Sequence 140, App
105	11	2.0	3270	7	US-11-063-439-65	Sequence 65, Appl	178	11	2.0	3503	7	US-11-063-439-214	Sequence 214, App
106	11	2.0	3317	7	US-11-063-439-157	Sequence 157, Appl	179	11	2.0	3503	7	US-11-063-439-225	Sequence 225, App
107	11	2.0	3397	7	US-11-063-439-245	Sequence 245, App	180	11	2.0	3504	7	US-11-063-439-131	Sequence 131, App
108	11	2.0	3445	7	US-11-063-439-242	Sequence 242, App	181	11	2.0	3504	7	US-11-063-439-119	Sequence 119, App
109	11	2.0	3457	7	US-11-063-439-194	Sequence 194, App	182	11	2.0	3504	7	US-11-063-439-148	Sequence 148, App
110	11	2.0	3461	7	US-11-063-439-193	Sequence 193, App	183	11	2.0	3505	7	US-11-063-439-18	Sequence 18, Appl
111	11	2.0	3461	7	US-11-063-439-205	Sequence 205, App	184	11	2.0	3506	7	US-11-063-439-105	Sequence 105, App
112	11	2.0	3462	7	US-11-063-439-217	Sequence 217, App	185	11	2.0	3506	7	US-11-063-439-232	Sequence 232, App
113	11	2.0	3463	7	US-11-063-439-250	Sequence 250, App	186	11	2.0	3507	7	US-11-063-439-5	Sequence 5, Appl1
114	11	2.0	3465	7	US-11-063-439-89	Sequence 89, Appl	187	11	2.0	3507	7	US-11-063-439-44	Sequence 44, Appl
115	11	2.0	3466	7	US-11-063-439-169	Sequence 169, App	188	11	2.0	3507	7	US-11-063-439-53	Sequence 53, Appl
116	11	2.0	3467	7	US-11-063-439-253	Sequence 253, App	189	11	2.0	3507	7	US-11-063-439-58	Sequence 58, Appl
117	11	2.0	3467	7	US-11-063-439-71	Sequence 71, Appl	190	11	2.0	3507	7	US-11-063-439-70	Sequence 70, Appl
118	11	2.0	3467	7	US-11-063-439-127	Sequence 127, App	191	11	2.0	3507	7	US-11-063-439-215	Sequence 215, App
119	11	2.0	3471	7	US-11-063-439-256	Sequence 256, App	192	11	2.0	3508	7	US-11-063-439-39	Sequence 39, Appl
120	11	2.0	3472	7	US-11-063-439-243	Sequence 243, App	193	11	2.0	3508	7	US-11-063-439-55	Sequence 55, Appl
121	11	2.0	3475	7	US-11-063-439-67	Sequence 67, Appl	194	11	2.0	3508	7	US-11-063-439-60	Sequence 60, Appl
122	11	2.0	3475	7	US-11-063-439-210	Sequence 210, App	195	11	2.0	3508	7	US-11-063-439-136	Sequence 136, App
123	11	2.0	3475	7	US-11-063-439-251	Sequence 251, App	196	11	2.0	3509	7	US-11-063-439-234	Sequence 234, App
124	11	2.0	3475	7	US-11-063-439-252	Sequence 252, App	197	11	2.0	3510	7	US-11-063-439-176	Sequence 176, App
125	11	2.0	3476	7	US-11-063-439-211	Sequence 211, App	198	11	2.0	3510	7	US-11-063-439-223	Sequence 223, App
126	11	2.0	3478	7	US-11-063-439-220	Sequence 220, App	199	11	2.0	3510	7	US-11-063-439-241	Sequence 241, App
127	11	2.0	3479	7	US-11-063-439-135	Sequence 135, App	200	11	2.0	3511	7	US-11-063-439-26	Sequence 26, Appl
128	11	2.0	3479	7	US-11-063-439-255	Sequence 255, App	201	11	2.0	3511	7	US-11-063-439-29	Sequence 29, Appl
129	11	2.0	3481	7	US-11-063-439-185	Sequence 185, App	202	11	2.0	3511	7	US-11-063-439-190	Sequence 190, App
130	11	2.0	3483	7	US-11-063-439-170	Sequence 170, App	203	11	2.0	3511	7	US-11-063-439-201	Sequence 201, App
131	11	2.0	3486	7	US-11-063-439-195	Sequence 195, App	204	11	2.0	3511	7	US-11-063-439-203	Sequence 203, App
132	11	2.0	3486	7	US-11-063-439-216	Sequence 216, App	205	11	2.0	3511	7	US-11-063-439-224	Sequence 224, App
133	11	2.0	3486	7	US-11-063-439-258	Sequence 258, App	206	11	2.0	3512	7	US-11-063-439-34	Sequence 34, Appl
134	11	2.0	3487	7	US-11-063-439-162	Sequence 162, App	207	11	2.0	3512	7	US-11-063-439-49	Sequence 49, Appl
135	11	2.0	3488	7	US-11-063-439-153	Sequence 153, App	208	11	2.0	3512	7	US-11-063-439-72	Sequence 72, Appl
136	11	2.0	3488	7	US-11-063-439-174	Sequence 174, App	209	11	2.0	3512	7	US-11-063-439-83	Sequence 83, Appl
137	11	2.0	3488	7	US-11-063-439-200	Sequence 200, App	210	11	2.0	3512	7	US-11-063-439-132	Sequence 132, App
138	11	2.0	3488	7	US-11-063-439-240	Sequence 240, App	211	11	2.0	3512	7	US-11-063-439-191	Sequence 191, App
139	11	2.0	3489	7	US-11-063-439-257	Sequence 257, App	212	11	2.0	3513	7	US-11-063-439-99	Sequence 99, Appl
140	11	2.0	3490	7	US-11-063-439-219	Sequence 219, App	213	11	2.0	3513	7	US-11-063-439-204	Sequence 204, Appl
141	11	2.0	3490	7	US-11-063-439-114	Sequence 114, App	214	11	2.0	3514	7	US-11-063-439-46	Sequence 46, Appl
142	11	2.0	3490	7	US-11-063-439-118	Sequence 118, App	215	11	2.0	3516	7	US-11-063-439-161	Sequence 161, App
143	11	2.0	3490	7	US-11-063-439-244	Sequence 244, App	216	11	2.0	3518	7	US-11-063-439-64	Sequence 64, Appl
144	11	2.0	3491	7	US-11-063-439-121	Sequence 121, App	217	11	2.0	3518	7	US-11-063-439-154	Sequence 154, App
145	11	2.0	3491	7	US-11-063-439-207	Sequence 207, App	218	11	2.0	3518	7	US-11-063-439-287	Sequence 287, App
146	11	2.0	3491	7	US-11-063-439-228	Sequence 228, App	219	11	2.0	3519	7	US-11-063-439-85	Sequence 85, Appl
147	11	2.0	3493	7	US-11-063-439-103	Sequence 103, App	220	11	2.0	3519	7	US-11-063-439-92	Sequence 92, Appl
148	11	2.0	3493	7	US-11-063-439-113	Sequence 113, App	221	11	2.0	3519	7	US-11-063-439-109	Sequence 109, App
149	11	2.0	3494	7	US-11-063-439-50	Sequence 50, Appl	222	11	2.0	3519	7	US-11-063-439-187	Sequence 187, App
150	11	2.0	3494	7	US-11-063-439-117	Sequence 117, App	223	11	2.0	3520	7	US-11-063-439-9	Sequence 9, Appl1
151	11	2.0	3494	7	US-11-063-439-235	Sequence 235, App	224	11	2.0	3520	7	US-11-063-439-112	Sequence 112, App
152	11	2.0	3494	7	US-11-063-439-239	Sequence 239, App	225	11	2.0	3520	7	US-11-063-439-202	Sequence 202, App
153	11	2.0	3495	7	US-11-063-439-141	Sequence 141, App	226	11	2.0	3521	7	US-11-063-439-63	Sequence 63, Appl
154	11	2.0	3495	7	US-11-063-439-181	Sequence 181, App	227	11	2.0	3522	7	US-11-063-439-80	Sequence 80, App
155	11	2.0	3495	7	US-11-063-439-198	Sequence 198, App	228	11	2.0	3523	7	US-11-063-439-146	Sequence 146, App
156	11	2.0	3496	7	US-11-063-439-123	Sequence 123, App	229	11	2.0	3523	7	US-11-063-439-152	Sequence 152, App
157	11	2.0	3496	7	US-11-063-439-125	Sequence 125, App	230	11	2.0	3524	7	US-11-063-439-13	Sequence 13, Appl
158	11	2.0	3496	7	US-11-063-439-145	Sequence 145, App	231	11	2.0	3524	7	US-11-063-439-61	Sequence 61, Appl
159	11	2.0	3496	7	US-11-063-439-151	Sequence 151, App	232	11	2.0	3524	7	US-11-063-439-124	Sequence 124, App
160	11	2.0	3496	7	US-11-063-439-237	Sequence 237, App	233	11	2.0	3524	7	US-11-063-439-179	Sequence 179, App
161	11	2.0	3497	7	US-11-063-439-133	Sequence 133, App	234	11	2.0	3525	7	US-11-063-439-11	Sequence 11, Appl
162	11	2.0	3498	7	US-11-063-439-27	Sequence 27, Appl	235	11	2.0	3526	7	US-11-063-439-36	Sequence 36, Appl
163	11	2.0	3498	7	US-11-063-439-142	Sequence 142, App	236	11	2.0	3526	7	US-11-063-439-130	Sequence 130, Appl
164	11	2.0	3498	7	US-11-063-439-164	Sequence 164, App	237	11	2.0	3528	7	US-11-063-439-32	Sequence 32, Appl
165	11	2.0	3498	7	US-11-063-439-209	Sequence 209, App	238	11	2.0	3528	7	US-11-063-439-38	Sequence 38, Appl
166	11	2.0	3498	7	US-11-063-439-238	Sequence 238, App	239	11	2.0	3528	7	US-11-063-439-155	Sequence 155, App
167	11	2.0	3499	7	US-11-063-439-96	Sequence 96, Appl	240	11	2.0	3528	7	US-11-063-439-156	Sequence 156, App
168	11	2.0	3499	7	US-11-063-439-222	Sequence 222, App	241	11	2.0	3529	7	US-11-063-439-16	Sequence 16, Appl
169	11	2.0	3499	7	US-11-063-439-246	Sequence 246, App	242	11	2.0	3529	7	US-11-063-439-37	Sequence 37, Appl
170	11	2.0	3500	7	US-11-063-439-110	Sequence 110, App	243	11	2.0	3529	7	US-11-063-439-87	Sequence 87, Appl
171	11	2.0	3500	7	US-11-063-439-218	Sequence 218, App	244	11	2.0	3531	7	US-11-063-439-139	Sequence 139, App

245	11	2.0	3531	7	US-11-063-439-206	Sequence 206, App	318	10	1.8	1523	7	US-11-174-307B-786	Sequence 786, App
246	11	2.0	3531	7	US-11-063-439-286	Sequence 286, App	319	10	1.8	1540	7	US-11-174-307B-2428	Sequence 2428, App
247	11	2.0	3533	7	US-11-063-439-14	Sequence 14, App	320	10	1.8	1550	7	US-11-174-307B-2158	Sequence 2158, App
248	11	2.0	3534	7	US-11-063-439-132	Sequence 132, App	321	10	1.8	1575	7	US-11-174-307B-2648	Sequence 2648, App
249	11	2.0	3536	7	US-11-063-439-31	Sequence 31, App	322	10	1.8	1684	7	US-11-174-307B-402	Sequence 402, App
250	11	2.0	3537	7	US-11-063-439-7	Sequence 7, App	323	10	1.8	1684	7	US-11-174-307B-1398	Sequence 1398, App
251	11	2.0	3537	7	US-11-063-439-111	Sequence 11, App	324	10	1.8	1696	7	US-11-174-307B-1378	Sequence 1378, App
252	11	2.0	3537	7	US-11-063-439-231	Sequence 231, App	325	10	1.8	1647	7	US-11-174-307B-2852	Sequence 2852, App
253	11	2.0	3540	7	US-11-063-439-188	Sequence 188, App	326	10	1.8	1917	7	US-11-174-307B-2172	Sequence 2172, App
254	11	2.0	3541	7	US-11-063-439-229	Sequence 229, App	327	10	1.8	1933	7	US-11-174-307B-2226	Sequence 2226, App
255	11	2.0	3544	7	US-11-063-439-158	Sequence 158, App	328	10	1.8	1942	7	US-11-174-307B-1096	Sequence 1096, App
256	11	2.0	3544	7	US-11-063-439-159	Sequence 159, App	329	10	1.8	2016	7	US-11-174-307B-1352	Sequence 1352, App
257	11	2.0	3544	7	US-11-063-439-197	Sequence 197, App	330	10	1.8	2247	7	US-11-174-307B-5510	Sequence 5510, App
258	11	2.0	3551	7	US-11-063-439-138	Sequence 138, App	331	10	1.8	2328	7	US-11-174-307B-1336	Sequence 1336, App
259	11	2.0	3561	7	US-11-063-439-120	Sequence 120, App	332	10	1.8	3392	7	US-11-174-307B-1854	Sequence 1854, App
260	11	2.0	3568	7	US-11-063-439-165	Sequence 165, App	333	9	1.6	93	6	US-10-953-349-35941	Sequence 35941, A
261	11	2.0	3573	7	US-11-063-439-150	Sequence 150, App	334	9	1.6	100	6	US-10-953-349-35940	Sequence 35940, A
262	11	2.0	3575	7	US-11-063-439-279	Sequence 279, App	335	9	1.6	101	6	US-10-953-349-35939	Sequence 35939, A
263	11	2.0	3578	7	US-11-063-439-74	Sequence 74, App	336	9	1.6	142	6	US-10-953-349-7026	Sequence 7026, App
264	11	2.0	3587	7	US-11-063-439-260	Sequence 260, App	337	9	1.6	176	6	US-10-953-349-7025	Sequence 7025, App
265	11	2.0	3588	7	US-11-063-439-52	Sequence 52, App	338	9	1.6	224	6	US-10-953-349-7024	Sequence 7024, App
266	11	2.0	3595	7	US-11-063-439-264	Sequence 264, App	339	9	1.6	287	7	US-11-056-355B-83065	Sequence 83065, A
267	11	2.0	3595	7	US-11-063-439-269	Sequence 269, App	340	9	1.6	313	6	US-10-449-902-45378	Sequence 45378, A
268	11	2.0	3599	7	US-11-063-439-271	Sequence 271, App	341	9	1.6	315	6	US-10-449-902-54725	Sequence 54725, A
269	11	2.0	3600	7	US-11-063-439-262	Sequence 262, App	342	9	1.6	315	6	US-10-449-902-54728	Sequence 54728, A
270	11	2.0	3601	7	US-11-063-439-266	Sequence 266, App	343	9	1.6	315	6	US-10-953-349-17978	Sequence 17978, A
271	11	2.0	3603	7	US-11-063-439-289	Sequence 289, App	344	9	1.6	324	6	US-11-056-355B-58249	Sequence 58249, A
272	11	2.0	3617	7	US-11-063-439-284	Sequence 284, App	345	9	1.6	324	6	US-10-953-349-17977	Sequence 17977, A
273	11	2.0	3650	7	US-11-063-439-280	Sequence 280, App	346	9	1.6	352	7	US-11-056-355B-58248	Sequence 58248, A
274	11	2.0	3651	7	US-11-063-439-268	Sequence 268, App	347	9	1.6	366	6	US-11-056-355B-68835	Sequence 68835, A
275	11	2.0	3657	7	US-11-063-439-129	Sequence 129, App	348	9	1.6	463	7	US-10-449-902-334885	Sequence 34885, A
276	11	2.0	3660	7	US-11-063-439-267	Sequence 267, App	349	9	1.6	490	6	US-11-174-307B-80	Sequence 80, App
277	11	2.0	3661	7	US-11-063-439-277	Sequence 277, App	350	9	1.6	490	6	US-10-449-902-28793	Sequence 28793, A
278	11	2.0	3661	7	US-11-063-439-281	Sequence 281, App	351	9	1.6	490	6	US-10-449-902-50419	Sequence 50419, A
279	11	2.0	3700	7	US-11-063-439-290	Sequence 290, App	352	9	1.6	586	6	US-10-449-902-37943	Sequence 37943, A
280	11	2.0	3711	7	US-11-063-439-261	Sequence 261, App	353	9	1.6	845	7	US-11-330-403-16179	Sequence 16179, A
281	11	2.0	3729	7	US-11-063-439-274	Sequence 274, App	354	9	1.6	857	7	US-11-174-307B-1232	Sequence 1232, App
282	11	2.0	3741	7	US-11-063-439-265	Sequence 265, App	355	9	1.6	1022	7	US-11-174-307B-4710	Sequence 4710, App
283	11	2.0	3742	7	US-11-063-439-283	Sequence 283, App	356	9	1.6	1065	7	US-11-174-307B-370	Sequence 370, App
284	11	2.0	3769	7	US-11-063-439-285	Sequence 285, App	357	9	1.6	1098	7	US-11-174-307B-2006	Sequence 2006, App
285	10	1.8	29	7	US-11-134-871-1223	Sequence 1223, App	358	9	1.6	1162	7	US-11-174-307B-2244	Sequence 2244, App
286	10	1.8	127	6	US-10-953-349-410	Sequence 410, App	359	9	1.6	1199	7	US-11-174-307B-2390	Sequence 2390, App
287	10	1.8	127	6	US-10-953-349-10505	Sequence 10505, A	360	9	1.6	1293	6	US-10-449-902-41380	Sequence 41380, A
288	10	1.8	145	6	US-10-953-349-409	Sequence 409, App	361	9	1.6	1302	7	US-11-174-307B-4424	Sequence 4424, App
289	10	1.8	217	6	US-10-953-349-10834	Sequence 10834, A	362	9	1.6	1313	7	US-11-174-307B-2312	Sequence 2312, App
290	10	1.8	222	7	US-11-056-355B-59365	Sequence 59365, A	363	9	1.6	1318	7	US-11-174-307B-1814	Sequence 1818, App
291	10	1.8	259	6	US-10-953-349-10833	Sequence 10833, A	364	9	1.6	1336	7	US-11-174-307B-532	Sequence 532, App
292	10	1.8	320	7	US-10-953-349-10832	Sequence 10832, A	365	9	1.6	1413	7	US-11-174-307B-936	Sequence 936, App
293	10	1.8	320	7	US-11-056-355B-41767	Sequence 41767, A	366	9	1.6	1456	7	US-11-174-307B-1164	Sequence 1164, App
294	10	1.8	335	6	US-10-374-780A-2066	Sequence 2066, App	367	9	1.6	1459	7	US-11-174-307B-1350	Sequence 1350, App
295	10	1.8	340	6	US-10-953-349-24393	Sequence 24393, A	368	9	1.6	1460	7	US-11-174-307B-5280	Sequence 5280, App
296	10	1.8	394	7	US-11-056-355B-41542	Sequence 41542, A	369	9	1.6	1462	7	US-11-174-307B-1056	Sequence 1056, App
297	10	1.8	395	7	US-11-056-355B-41541	Sequence 41541, A	370	9	1.6	1469	7	US-11-174-307B-1890	Sequence 1890, App
298	10	1.8	410	7	US-11-056-355B-41540	Sequence 41540, A	371	9	1.6	1470	7	US-11-174-307B-1280	Sequence 1280, App
299	10	1.8	427	6	US-10-953-349-24392	Sequence 24392, A	372	9	1.6	1509	7	US-11-174-307B-1838	Sequence 1838, App
300	10	1.8	445	6	US-10-953-349-24391	Sequence 24391, A	373	9	1.6	1510	7	US-11-174-307B-538	Sequence 538, App
301	10	1.8	660	7	US-11-174-307B-408	Sequence 408, App	374	9	1.6	1577	7	US-11-174-307B-396	Sequence 396, App
302	10	1.8	1005	7	US-11-174-307B-1158	Sequence 1158, App	375	9	1.6	1598	7	US-11-174-307B-5072	Sequence 5072, App
303	10	1.8	1042	7	US-11-174-307B-356	Sequence 356, App	376	9	1.6	1600	7	US-11-174-307B-936	Sequence 936, App
304	10	1.8	1050	7	US-11-174-307B-724	Sequence 724, App	377	9	1.6	1629	7	US-11-174-307B-2828	Sequence 2828, App
305	10	1.8	1061	7	US-11-174-307B-328	Sequence 328, App	378	9	1.6	1674	7	US-11-174-307B-1328	Sequence 1328, App
306	10	1.8	1192	7	US-11-174-307B-394	Sequence 394, App	379	9	1.6	1686	7	US-11-174-307B-1802	Sequence 1802, App
307	10	1.8	1192	7	US-11-174-307B-1832	Sequence 1832, App	380	9	1.6	1693	7	US-11-174-307B-1910	Sequence 1910, App
308	10	1.8	1220	7	US-11-174-307B-1542	Sequence 1542, App	381	9	1.6	1696	7	US-11-174-307B-1464	Sequence 1464, App
309	10	1.8	1298	7	US-11-174-307B-2774	Sequence 2774, App	382	9	1.6	1732	7	US-11-174-307B-1086	Sequence 1086, App
310	10	1.8	1306	7	US-11-174-307B-1254	Sequence 1254, App	383	9	1.6	1739	7	US-11-174-307B-3342	Sequence 3342, App
311	10	1.8	1348	7	US-11-174-307B-2882	Sequence 2882, App	384	9	1.6	1771	7	US-11-174-307B-928	Sequence 928, App
312	10	1.8	1355	7	US-11-174-307B-2278	Sequence 2278, App	385	9	1.6	1775	7	US-11-174-307B-526	Sequence 526, App
313	10	1.8	1363	7	US-11-174-307B-358	Sequence 358, App	386	9	1.6	1806	7	US-11-174-307B-972	Sequence 972, App
314	10	1.8	1363	7	US-11-174-307B-2430	Sequence 2430, App	387	9	1.6	1811	7	US-11-174-307B-924	Sequence 924, App
315	10	1.8	1457	7	US-11-174-307B-1094	Sequence 1094, App	388	9	1.6	1818	7	US-11-174-307B-586	Sequence 586, App
316	10	1.8	1482	7	US-11-174-307B-880	Sequence 880, App	389	9	1.6	1851	7	US-11-174-307B-2240	Sequence 2240, App
317	10	1.8	1484	7	US-11-174-307B-424	Sequence 424, App	390	9	1.6	1919	7	US-11-174-307B-578	Sequence 578, App

391	9	1.6	1925	7	US-11-174-307B-2362	Sequence 2362, Ap
392	9	1.6	1951	7	US-11-174-307B-1118	Sequence 3118, Ap
393	9	1.6	1957	7	US-11-174-307B-2230	Sequence 2230, Ap
394	9	1.6	1962	7	US-11-174-307B-1446	Sequence 1446, Ap
395	9	1.6	2003	7	US-11-174-307B-5496	Sequence 5496, Ap
396	9	1.6	2062	7	US-11-174-307B-1454	Sequence 1454, Ap
397	9	1.6	2143	7	US-11-174-307B-2118	Sequence 2118, Ap
398	9	1.6	2142	7	US-11-174-307B-2550	Sequence 2550, Ap
399	9	1.6	4243	7	US-11-174-307B-1722	Sequence 1722, Ap
400	8	1.4	25	7	US-11-316-535A-54	Sequence 54, Appl
401	8	1.4	135	6	US-10-449-902-29445	Sequence 29445, A
402	8	1.4	179	7	US-11-330-403-18758	Sequence 18758, A
403	8	1.4	198	7	US-11-056-355B-28568	Sequence 28568, A
404	8	1.4	198	7	US-11-056-355B-2158	Sequence 32158, A
405	8	1.4	202	7	US-11-056-355B-2157	Sequence 28567, A
406	8	1.4	202	7	US-11-056-355B-22157	Sequence 32157, A
407	8	1.4	204	7	US-11-056-355B-28601	Sequence 28601, A
408	8	1.4	204	7	US-11-056-355B-32191	Sequence 32191, A
409	8	1.4	204	7	US-11-056-355B-50009	Sequence 50009, A
410	8	1.4	204	7	US-11-056-355B-71349	Sequence 71349, A
411	8	1.4	204	7	US-11-056-355B-98417	Sequence 98417, A
412	8	1.4	204	7	US-11-056-355B-109656	Sequence 109656, A
413	8	1.4	210	6	US-10-374-780A-636	Sequence 636, App
414	8	1.4	232	7	US-11-056-355B-50008	Sequence 50008, A
415	8	1.4	232	7	US-11-056-355B-71348	Sequence 71348, A
416	8	1.4	232	7	US-11-056-355B-98416	Sequence 98416, A
417	8	1.4	232	7	US-11-056-355B-109655	Sequence 109655, A
418	8	1.4	246	6	US-10-449-902-30278	Sequence 30278, A
419	8	1.4	246	6	US-10-449-902-55271	Sequence 55271, A
420	8	1.4	254	6	US-10-566-644-332	Sequence 332, App
421	8	1.4	297	7	US-11-056-355B-19575	Sequence 19575, A
422	8	1.4	307	7	US-11-056-355B-80070	Sequence 80070, A
423	8	1.4	311	6	US-10-449-902-35626	Sequence 35626, A
424	8	1.4	311	7	US-11-056-355B-80069	Sequence 80069, A
425	8	1.4	318	7	US-11-056-355B-101841	Sequence 101841, A
426	8	1.4	318	7	US-11-056-355B-113080	Sequence 113080, A
427	8	1.4	323	6	US-10-374-780A-2178	Sequence 2178, Ap
428	8	1.4	323	7	US-11-056-355B-80068	Sequence 80068, A
429	8	1.4	337	6	US-10-449-902-31780	Sequence 31780, A
430	8	1.4	357	7	US-11-056-355B-98415	Sequence 98415, A
431	8	1.4	357	7	US-11-056-355B-109654	Sequence 109654, A
432	8	1.4	358	6	US-10-449-902-35482	Sequence 35482, A
433	8	1.4	359	7	US-11-056-355B-50007	Sequence 50007, A
434	8	1.4	359	7	US-11-056-355B-71347	Sequence 71347, A
435	8	1.4	360	7	US-11-056-355B-67932	Sequence 67932, A
436	8	1.4	368	6	US-10-449-902-38075	Sequence 38075, A
437	8	1.4	376	7	US-11-056-355B-16351	Sequence 16351, A
438	8	1.4	381	7	US-11-330-403-15280	Sequence 15280, A
439	8	1.4	389	7	US-11-056-355B-67931	Sequence 67931, A
440	8	1.4	399	6	US-10-449-902-41963	Sequence 41963, A
441	8	1.4	411	6	US-10-449-902-39672	Sequence 39672, A
442	8	1.4	412	6	US-10-953-349-22109	Sequence 22109, A
443	8	1.4	412	7	US-11-056-355B-57200	Sequence 57200, A
444	8	1.4	416	6	US-10-449-902-42045	Sequence 42045, A
445	8	1.4	420	6	US-10-449-902-41169	Sequence 41169, A
446	8	1.4	420	6	US-11-056-355B-67930	Sequence 67930, A
447	8	1.4	422	6	US-10-449-902-36574	Sequence 36574, A
448	8	1.4	424	6	US-10-449-902-40147	Sequence 40147, A
449	8	1.4	430	6	US-10-953-349-22108	Sequence 22108, A
450	8	1.4	430	7	US-11-056-355B-57199	Sequence 57199, A
451	8	1.4	434	7	US-11-056-355B-17176	Sequence 17176, A
452	8	1.4	434	7	US-11-056-355B-86255	Sequence 86255, A
453	8	1.4	439	7	US-11-056-355B-49440	Sequence 49440, A
454	8	1.4	451	7	US-11-056-355B-17175	Sequence 17175, A
455	8	1.4	456	7	US-11-056-355B-49439	Sequence 49439, A
456	8	1.4	464	6	US-10-953-349-22107	Sequence 22107, A
457	8	1.4	464	7	US-11-056-355B-57198	Sequence 57198, A
458	8	1.4	466	7	US-11-056-355B-49438	Sequence 49438, A
459	8	1.4	470	6	US-11-056-355B-17174	Sequence 17174, A
460	8	1.4	484	6	US-10-449-902-44153	Sequence 44153, A
461	8	1.4	484	6	US-10-449-902-45190	Sequence 45190, A
462	8	1.4	505	7	US-11-330-403-611	Sequence 611, App
463	8	1.4	513	7	US-11-330-403-11394	Sequence 11394, A
464	8	1.4	537	6	US-10-449-902-38511	Sequence 38511, A
465	8	1.4	558	6	US-10-953-349-10659	Sequence 10659, A
466	8	1.4	558	7	US-11-056-355B-45019	Sequence 45019, A
467	8	1.4	558	7	US-11-056-355B-47210	Sequence 47210, A
468	8	1.4	558	7	US-11-056-355B-19889	Sequence 49869, Ap
469	8	1.4	558	7	US-11-056-355B-108349	Sequence 108349, A
470	8	1.4	558	7	US-11-056-355B-108588	Sequence 108588, A
471	8	1.4	558	7	US-11-056-355B-108596	Sequence 108596, A
472	8	1.4	558	7	US-11-056-355B-119828	Sequence 119828, A
473	8	1.4	558	7	US-11-056-355B-119827	Sequence 119827, A
474	8	1.4	558	7	US-11-056-355B-119835	Sequence 119835, A
475	8	1.4	559	6	US-10-953-349-10658	Sequence 10658, A
476	8	1.4	559	7	US-11-056-355B-50018	Sequence 45018, A
477	8	1.4	559	7	US-11-056-355B-47149	Sequence 47149, A
478	8	1.4	559	7	US-11-056-355B-49868	Sequence 49868, A
479	8	1.4	559	7	US-11-056-355B-108348	Sequence 108348, A
480	8	1.4	559	7	US-11-056-355B-108587	Sequence 108587, A
481	8	1.4	559	7	US-11-056-355B-108595	Sequence 108595, A
482	8	1.4	559	7	US-11-056-355B-119587	Sequence 119587, A
483	8	1.4	559	7	US-11-056-355B-119826	Sequence 119826, A
484	8	1.4	559	7	US-11-056-355B-119834	Sequence 119834, A
485	8	1.4	572	7	US-11-174-307B-5524	Sequence 5524, Ap
486	8	1.4	577	6	US-10-449-902-50663	Sequence 50663, A
487	8	1.4	587	6	US-10-953-349-10657	Sequence 10657, A
488	8	1.4	587	7	US-11-056-355B-45017	Sequence 45017, A
489	8	1.4	587	7	US-11-056-355B-47148	Sequence 47148, A
490	8	1.4	587	7	US-11-056-355B-49867	Sequence 49867, A
491	8	1.4	587	7	US-11-056-355B-108347	Sequence 108347, A
492	8	1.4	587	7	US-11-056-355B-108586	Sequence 108586, A
493	8	1.4	587	7	US-11-056-355B-108594	Sequence 108594, A
494	8	1.4	587	7	US-11-056-355B-119586	Sequence 119586, A
495	8	1.4	587	7	US-11-056-355B-119825	Sequence 119825, A
496	8	1.4	587	7	US-11-056-355B-119833	Sequence 119833, A
497	8	1.4	592	6	US-10-566-741-54	Sequence 54, Appl
498	8	1.4	595	7	US-11-330-403-6027	Sequence 6027, Ap
499	8	1.4	613	7	US-11-174-307B-1659	Sequence 1659, Ap
500	8	1.4	644	7	US-11-174-307B-4730	Sequence 4730, Ap
501	8	1.4	658	6	US-10-449-902-54667	Sequence 54667, A
502	8	1.4	678	7	US-11-174-307B-5924	Sequence 5924, Ap
503	8	1.4	785	7	US-11-174-307B-3002	Sequence 3002, Ap
504	8	1.4	889	7	US-11-174-307B-390	Sequence 390, App
505	8	1.4	893	6	US-10-449-902-52708	Sequence 52708, A
506	8	1.4	929	7	US-11-174-307B-8296	Sequence 8296, Ap
507	8	1.4	994	7	US-11-174-307B-918	Sequence 918, App
508	8	1.4	1015	7	US-11-174-307B-330	Sequence 330, App
509	8	1.4	1017	7	US-11-174-307B-956	Sequence 956, App
510	8	1.4	1057	7	US-11-174-307B-2602	Sequence 2602, Ap
511	8	1.4	1060	7	US-11-174-307B-2216	Sequence 2216, Ap
512	8	1.4	1084	7	US-11-174-307B-5106	Sequence 5106, Ap
513	8	1.4	1119	7	US-11-174-307B-456	Sequence 456, App
514	8	1.4	1144	7	US-11-174-307B-2720	Sequence 2720, Ap
515	8	1.4	1220	7	US-11-174-307B-3330	Sequence 3330, Ap
516	8	1.4	1248	7	US-11-174-307B-1262	Sequence 1262, Ap
517	8	1.4	1248	7	US-11-174-307B-2506	Sequence 2506, Ap
518	8	1.4	1286	7	US-11-174-307B-1334	Sequence 1334, Ap
519	8	1.4	1287	7	US-11-174-307B-1166	Sequence 1166, Ap
520	8	1.4	1303	7	US-11-174-307B-1330	Sequence 1330, App
521	8	1.4	1303	7	US-11-174-307B-1570	Sequence 1570, Ap
522	8	1.4	1307	7	US-11-174-307B-4164	Sequence 4164, Ap
523	8	1.4	1315	7	US-11-174-307B-8248	Sequence 8248, Ap
524	8	1.4	1323	7	US-11-174-307B-1734	Sequence 1734, Ap
525	8	1.4	1325	7	US-11-174-307B-450	Sequence 450, App
526	8	1.4	1357	7	US-11-174-307B-1724	Sequence 1724, Ap
527	8	1.4	1359	7	US-11-174-307B-974	Sequence 974, App
528	8	1.4	1415	7	US-11-174-307B-1762	Sequence 1762, App
529	8	1.4	1422	7	US-11-063-439-227	Sequence 227, App
530	8	1.4	1429	7	US-11-174-307B-2274	Sequence 2274, Ap
531	8	1.4	1431	7	US-11-174-307B-1000	Sequence 1000, Ap
532	8	1.4	1431	7	US-11-174-307B-1112	Sequence 1112, Ap
533	8	1.4	1435	6	US-10-196-749-581	Sequence 581, App
534	8	1.4	1443	6	US-11-174-307B-2906	Sequence 2906, App
535	8	1.4	1449	7	US-11-174-307B-134	Sequence 134, App
536	8	1.4	1449	7	US-11-174-307B-336	Sequence 336, App


```

537 8 1.4 1485 7 US-11-174-307B-2206 Sequence 2206, Ap
538 8 1.4 1488 7 US-11-174-307B-4372 Sequence 4372, Ap
539 8 1.4 1484 7 US-11-174-307B-820 Sequence 820, App
540 8 1.4 1485 7 US-11-174-307B-1656 Sequence 1656, Ap
541 8 1.4 1485 7 US-11-174-307B-1154 Sequence 1154, Ap
542 8 1.4 1533 7 US-11-330-403-4077 Sequence 4077, Ap
543 8 1.4 1534 7 US-11-174-307B-2156 Sequence 2156, Ap
544 8 1.4 1540 7 US-11-174-307B-2418 Sequence 2418, Ap
545 8 1.4 1548 7 US-11-063-439-94 Sequence 94, Appl
546 8 1.4 1553 7 US-11-174-307B-380 Sequence 380, Appl
547 8 1.4 1602 7 US-11-174-307B-688 Sequence 688, App
548 8 1.4 1639 7 US-11-174-307B-1134 Sequence 1134, Ap
549 8 1.4 1652 7 US-11-174-307B-1188 Sequence 1188, Ap
550 8 1.4 1660 7 US-11-174-307B-2908 Sequence 2908, Ap
551 8 1.4 1685 7 US-11-174-307B-2212 Sequence 2212, Ap
552 8 1.4 1706 7 US-11-174-307B-2322 Sequence 2322, Ap
553 8 1.4 1713 7 US-11-174-307B-4894 Sequence 4894, Ap
554 8 1.4 1722 7 US-11-174-307B-350 Sequence 350, App
555 8 1.4 1745 7 US-11-174-307B-960 Sequence 960, App
556 8 1.4 1751 7 US-11-174-307B-1128 Sequence 1128, Ap
557 8 1.4 1768 7 US-11-174-307B-1136 Sequence 1136, Ap
558 8 1.4 1782 7 US-11-174-307B-1652 Sequence 1652, Ap
559 8 1.4 1791 7 US-11-174-307B-2026 Sequence 2026, Ap
560 8 1.4 1819 7 US-11-174-307B-822 Sequence 822, Ap
561 8 1.4 1831 7 US-11-174-307B-1034 Sequence 1034, Ap
562 8 1.4 1863 7 US-11-174-307B-988 Sequence 988, App
563 8 1.4 1882 7 US-11-174-307B-1388 Sequence 1388, Ap
564 8 1.4 1910 7 US-11-174-307B-2256 Sequence 2256, Ap
565 8 1.4 1913 7 US-11-063-439-163 Sequence 163, App
566 8 1.4 1913 7 US-11-174-307B-4028 Sequence 4028, Ap
567 8 1.4 1942 7 US-11-174-307B-1076 Sequence 1076, Ap
568 8 1.4 1945 7 US-11-174-307B-1972 Sequence 1972, Ap
569 8 1.4 1988 7 US-11-174-307B-2860 Sequence 2860, Ap
570 8 1.4 2105 7 US-11-174-307B-1554 Sequence 1554, Ap
571 8 1.4 2127 7 US-11-174-307B-748 Sequence 748, App
572 8 1.4 2180 7 US-11-174-307B-530 Sequence 530, App
573 8 1.4 2366 7 US-11-174-307B-804 Sequence 804, App
574 8 1.4 2440 7 US-11-174-307B-1294 Sequence 1294, Ap
575 8 1.4 2616 7 US-11-174-307B-1680 Sequence 1680, Ap
576 8 1.4 2643 7 US-11-174-307B-1864 Sequence 1864, Ap
577 8 1.4 3203 7 US-11-063-439-171 Sequence 171, App
578 8 1.4 3222 7 US-11-063-439-137 Sequence 137, App
579 8 1.4 3331 7 US-11-174-307B-1574 Sequence 1574, Ap
580 8 1.4 3447 7 US-11-063-439-21 Sequence 21, Appl
581 8 1.4 3462 7 US-11-063-439-41 Sequence 41, Appl
582 8 1.4 3469 7 US-11-063-439-106 Sequence 106, App
583 8 1.4 3470 7 US-11-063-439-143 Sequence 143, App
584 8 1.4 3473 7 US-11-063-439-199 Sequence 199, App
585 8 1.4 3478 7 US-11-063-439-79 Sequence 79, Appl
586 8 1.4 3478 7 US-11-063-439-95 Sequence 95, Appl
587 8 1.4 3478 7 US-11-063-439-134 Sequence 134, App
588 8 1.4 3479 7 US-11-063-439-25 Sequence 25, Appl
589 8 1.4 3480 7 US-11-063-439-57 Sequence 57, Appl
590 8 1.4 3480 7 US-11-063-439-78 Sequence 78, Appl
591 8 1.4 3481 7 US-11-063-439-22 Sequence 22, Appl
592 8 1.4 3481 7 US-11-063-439-90 Sequence 90, Appl
593 8 1.4 3482 7 US-11-063-439-48 Sequence 48, Appl
594 8 1.4 3482 7 US-11-063-439-97 Sequence 97, Appl
595 8 1.4 3483 7 US-11-063-439-4 Sequence 4, Appl
596 8 1.4 3483 7 US-11-063-439-28 Sequence 28, Appl
597 8 1.4 3483 7 US-11-063-439-98 Sequence 98, Appl
598 8 1.4 3486 7 US-11-063-439-167 Sequence 167, App
599 8 1.4 3487 7 US-11-063-439-17 Sequence 17, Appl
600 8 1.4 3490 7 US-11-063-439-13 Sequence 13, Appl
601 8 1.4 3490 7 US-11-063-439-126 Sequence 126, App
602 8 1.4 3491 7 US-11-063-439-33 Sequence 33, Appl
603 8 1.4 3493 7 US-11-063-439-6 Sequence 6, Appl
604 8 1.4 3493 7 US-11-063-439-12 Sequence 12, Appl
605 8 1.4 3493 7 US-11-063-439-35 Sequence 35, Appl
606 8 1.4 3493 7 US-11-063-439-47 Sequence 47, Appl
607 8 1.4 3494 7 US-11-063-439-1 Sequence 1, Appl
608 8 1.4 3494 7 US-11-063-439-104 Sequence 104, App
609 8 1.4 3496 7 US-11-063-439-173 Sequence 173, App

```

```

610 8 1.4 3496 7 US-11-063-439-230 Sequence 230, App
611 8 1.4 3497 7 US-11-063-439-102 Sequence 102, App
612 8 1.4 3497 7 US-11-063-439-172 Sequence 172, App
613 8 1.4 3497 7 US-11-063-439-180 Sequence 180, App
614 8 1.4 3499 7 US-11-063-439-116 Sequence 116, App
615 8 1.4 3499 7 US-11-063-439-147 Sequence 147, App
616 8 1.4 3500 7 US-11-063-439-84 Sequence 84, Appl
617 8 1.4 3502 7 US-11-063-439-221 Sequence 221, Appl
618 8 1.4 3506 7 US-11-063-439-77 Sequence 77, Appl
619 8 1.4 3507 7 US-11-063-439-196 Sequence 196, App
620 8 1.4 3508 7 US-11-063-439-166 Sequence 166, App
621 8 1.4 3508 7 US-11-063-439-168 Sequence 168, App
622 8 1.4 3509 7 US-11-063-439-178 Sequence 178, App
623 8 1.4 3511 7 US-11-063-439-130 Sequence 130, App
624 8 1.4 3515 7 US-11-063-439-101 Sequence 101, App
625 8 1.4 3517 7 US-11-063-439-8 Sequence 8, Appl
626 8 1.4 3518 7 US-11-063-439-107 Sequence 107, App
627 8 1.4 3524 7 US-11-063-439-93 Sequence 93, Appl
628 8 1.4 3528 7 US-11-063-439-182 Sequence 182, App
629 8 1.4 3541 7 US-11-063-439-128 Sequence 128, App
630 8 1.4 3542 7 US-11-063-439-212 Sequence 212, App
631 8 1.4 3552 7 US-11-063-439-192 Sequence 192, App
632 8 1.4 3605 7 US-11-063-439-213 Sequence 213, App
633 8 1.4 3974 7 US-11-063-439-276 Sequence 276, App
634 8 1.4 4440 6 US-10-196-749-525 Sequence 525, App
635 8 1.4 5179 7 US-11-105-233-185 Sequence 185, App

```

ALIGNMENTS

```

RESULT 1
US-10-471-571A-782
; Sequence 782, Application US/10471571A
; Publication No. US20060115490A1
GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026927WO
; CURRENT APPLICATION NUMBER: US/10/471,571A
; CURRENT FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: GB-0107661.1
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 5642
; SOFTWARE: SeqMin99, version 1.03
; SEQ ID NO 782
; LENGTH: 166
; TYPE: prt
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (1)..(166)
; OTHER INFORMATION: hypothetical protein
US-10-471-571A-782
Query Match 2.3%; Score 13; DB 6; Length 166;
Best Local Similarity 100.0%; Pred. No. 0.00018;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 392 OTTTTTTTTTT 404
Db 43 OTTTTTTTTTT 55

```

```

RESULT 2
US-10-953-349-34264
; Sequence 34264, Application US/10953349
; Publication No. US20060107345A1
GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: ENCODED THERAPY
; FILE REFERENCE: 2750-1579PUS2

```

```
/ CURRENT APPLICATION NUMBER: US/10/953,349
/ CURRENT FILING DATE: 2004-09-30
/ NUMBER OF SEQ ID NOS: 40252
/ SOFTWARE: PatentIn version 3.3
/ SEQ ID NO 34264
/ LENGTH: 264
/ TYPE: PRT
/ ORGANISM: Zea mays subsp. mays
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (156)..(158)
/ OTHER INFORMATION: Xaa can be any naturally occurring amino acid
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (165)..(165)
/ OTHER INFORMATION: Xaa can be any naturally occurring amino acid
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (171)..(171)
/ OTHER INFORMATION: Xaa can be any naturally occurring amino acid
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (177)..(177)
/ OTHER INFORMATION: Xaa can be any naturally occurring amino acid
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (186)..(186)
/ OTHER INFORMATION: Xaa can be any naturally occurring amino acid
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (204)..(204)
/ OTHER INFORMATION: Xaa can be any naturally occurring amino acid
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (215)..(215)
/ OTHER INFORMATION: Xaa can be any naturally occurring amino acid
/ US-10-953-349-34264
```

```
Query Match 2.2%; Score 12; DB 6; Length 264;
Best Local Similarity 100.0%; Pred. No. 0.0025;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy 393 TTTT TTTT TTTT TTTT TTTT 404
Db 13 TTTT TTTT TTTT TTTT TTTT 24

RESULT 3
US-10-953-349-11680
/ Sequence 11680, Application US/10953349
/ Publication No. US20060107345A1
/ GENERAL INFORMATION:
/ APPLICANT: ALEXANDROV, Nikolai et al.
/ TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
/ FILE REFERENCE: 2750-1579PUS2
/ CURRENT APPLICATION NUMBER: US/10/953,349
/ CURRENT FILING DATE: 2004-09-30
/ NUMBER OF SEQ ID NOS: 40252
/ SOFTWARE: PatentIn version 3.3
/ SEQ ID NO 11680
/ LENGTH: 659
/ TYPE: PRT
/ ORGANISM: Arabidopsis thaliana
/ US-10-953-349-11680
```

```
Query Match 2.2%; Score 12; DB 6; Length 659;
Best Local Similarity 100.0%; Pred. No. 0.0058;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy 393 TTTT TTTT TTTT TTTT TTTT 404
Db 53 TTTT TTTT TTTT TTTT TTTT 64
```

```
RESULT 4
US-11-063-439-66
/ Sequence 66, Application US/11063439
/ Publication No. US20060147371A1
/ GENERAL INFORMATION:
/ APPLICANT: TUSZYNSKI, JACK A.
/ APPLICANT: GREENWALD, HOWARD J.
/ APPLICANT: CURRY, STEPHEN H.
/ APPLICANT: GOSS, KENDRICK
/ TITLE OF INVENTION: WATER-SOLUBLE COMPOUND
/ FILE REFERENCE: 1034312-000027
/ CURRENT APPLICATION NUMBER: US/11/063,439
/ CURRENT FILING DATE: 2005-02-23
/ PRIOR APPLICATION NUMBER: 10/878,905
/ PRIOR FILING DATE: 2004-06-28
/ PRIOR APPLICATION NUMBER: 10/923,615
/ PRIOR FILING DATE: 2004-08-20
/ PRIOR APPLICATION NUMBER: 60/516,134
/ PRIOR FILING DATE: 2003-10-31
/ PRIOR APPLICATION NUMBER: 10/808,618
/ PRIOR FILING DATE: 2004-03-24
/ PRIOR APPLICATION NUMBER: 10/867,517
/ PRIOR FILING DATE: 2004-06-14
/ NUMBER OF SEQ ID NOS: 418
/ SOFTWARE: PatentIn version 3.3
/ SEQ ID NO 66
/ LENGTH: 1675
/ TYPE: PRT
/ ORGANISM: Pneumocystis carinii
/ US-11-063-439-66
```

```
Query Match 2.2%; Score 12; DB 7; Length 1675;
Best Local Similarity 100.0%; Pred. No. 0.014;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy 393 TTTT TTTT TTTT TTTT TTTT 404
Db 1382 TTTT TTTT TTTT TTTT TTTT 1393
```

```
RESULT 5
US-11-063-439-24
/ Sequence 24, Application US/11063439
/ Publication No. US20060147371A1
/ GENERAL INFORMATION:
/ APPLICANT: TUSZYNSKI, JACK A.
/ APPLICANT: GREENWALD, HOWARD J.
/ APPLICANT: CURRY, STEPHEN H.
/ APPLICANT: GOSS, KENDRICK
/ TITLE OF INVENTION: WATER-SOLUBLE COMPOUND
/ FILE REFERENCE: 1034312-000027
/ CURRENT APPLICATION NUMBER: US/11/063,439
/ CURRENT FILING DATE: 2005-02-23
/ PRIOR APPLICATION NUMBER: 10/878,905
/ PRIOR FILING DATE: 2004-06-28
/ PRIOR APPLICATION NUMBER: 10/923,615
/ PRIOR FILING DATE: 2004-08-20
/ PRIOR APPLICATION NUMBER: 60/516,134
/ PRIOR FILING DATE: 2003-10-31
/ PRIOR APPLICATION NUMBER: 10/808,618
/ PRIOR FILING DATE: 2004-03-24
/ PRIOR APPLICATION NUMBER: 10/867,517
/ PRIOR FILING DATE: 2004-06-14
/ NUMBER OF SEQ ID NOS: 418
/ SOFTWARE: PatentIn version 3.3
/ SEQ ID NO 24
/ LENGTH: 2823
/ TYPE: PRT
/ ORGANISM: Anemia phyllitidis
/ US-11-063-439-24
```

Query Match 2.2%; Score 12; DB 7; Length 2823;
Best Local Similarity 100.0%; Pred. No. 0.022;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 393 TTTT TTTT TTTT TTTT 404
DB 1189 TTTT TTTT TTTT TTTT 1200

RESULT 6

US-11-063-439-273
; Sequence 273, Application US/11063439
; Publication No. US20060147371A1
; GENERAL INFORMATION:
; APPLICANT: TUSZYNSKI, JACK A.
; APPLICANT: GREENWALD, HOWARD J.
; APPLICANT: CURRY, STEPHEN H.
; APPLICANT: GOSS, KENDRICK
; TITLE OF INVENTION: WATER-SOLUBLE COMPOUND
; FILE REFERENCE: 1034312-000027
; CURRENT APPLICATION NUMBER: US/11/063,439
; CURRENT FILING DATE: 2005-02-23
; PRIOR APPLICATION NUMBER: 10/878,905
; PRIOR FILING DATE: 2004-06-28
; PRIOR APPLICATION NUMBER: 10/923,615
; PRIOR FILING DATE: 2004-08-20
; PRIOR APPLICATION NUMBER: 60/516,134
; PRIOR FILING DATE: 2003-10-31
; PRIOR APPLICATION NUMBER: 10/808,618
; PRIOR FILING DATE: 2004-03-24
; PRIOR APPLICATION NUMBER: 10/667,517
; PRIOR FILING DATE: 2004-06-14
; NUMBER OF SEQ ID NOS: 418
; SOFTWARE: Patentin version 3.3
; SEQ ID NO 273
; LENGTH: 3342
; TYPE: PRT
; ORGANISM: Zea mays
US-11-063-439-273

Query Match 2.2%; Score 12; DB 7; Length 3342;
Best Local Similarity 100.0%; Pred. No. 0.025;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 392 TTTT TTTT TTTT TTTT 403
DB 2541 TTTT TTTT TTTT TTTT 2552

RESULT 7

US-11-063-439-68
; Sequence 68, Application US/11063439
; Publication No. US20060147371A1
; GENERAL INFORMATION:
; APPLICANT: TUSZYNSKI, JACK A.
; APPLICANT: GREENWALD, HOWARD J.
; APPLICANT: CURRY, STEPHEN H.
; APPLICANT: GOSS, KENDRICK
; TITLE OF INVENTION: WATER-SOLUBLE COMPOUND
; FILE REFERENCE: 1034312-000027
; CURRENT APPLICATION NUMBER: US/11/063,439
; CURRENT FILING DATE: 2005-02-23
; PRIOR APPLICATION NUMBER: 10/878,905
; PRIOR FILING DATE: 2004-06-28
; PRIOR APPLICATION NUMBER: 10/923,615
; PRIOR FILING DATE: 2004-08-20
; PRIOR APPLICATION NUMBER: 60/516,134
; PRIOR FILING DATE: 2003-10-31
; PRIOR APPLICATION NUMBER: 10/808,618
; PRIOR FILING DATE: 2004-03-24
; PRIOR APPLICATION NUMBER: 10/667,517
; PRIOR FILING DATE: 2004-06-14
; NUMBER OF SEQ ID NOS: 418

; SOFTWARE: Patentin version 3.3
; SEQ ID NO 68
; LENGTH: 3445
; TYPE: PRT
; ORGANISM: Avena sativa
US-11-063-439-68

Query Match 2.2%; Score 12; DB 7; Length 3445;
Best Local Similarity 100.0%; Pred. No. 0.026;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 393 TTTT TTTT TTTT TTTT 404
DB 1831 TTTT TTTT TTTT TTTT 1842

RESULT 8

US-11-063-439-248
; Sequence 248, Application US/11063439
; Publication No. US20060147371A1
; GENERAL INFORMATION:
; APPLICANT: TUSZYNSKI, JACK A.
; APPLICANT: GREENWALD, HOWARD J.
; APPLICANT: CURRY, STEPHEN H.
; APPLICANT: GOSS, KENDRICK
; TITLE OF INVENTION: WATER-SOLUBLE COMPOUND
; FILE REFERENCE: 1034312-000027
; CURRENT APPLICATION NUMBER: US/11/063,439
; CURRENT FILING DATE: 2005-02-23
; PRIOR APPLICATION NUMBER: 10/878,905
; PRIOR FILING DATE: 2004-06-28
; PRIOR APPLICATION NUMBER: 10/923,615
; PRIOR FILING DATE: 2004-08-20
; PRIOR APPLICATION NUMBER: 60/516,134
; PRIOR FILING DATE: 2003-10-31
; PRIOR APPLICATION NUMBER: 10/808,618
; PRIOR FILING DATE: 2004-03-24
; PRIOR APPLICATION NUMBER: 10/667,517
; PRIOR FILING DATE: 2004-06-14
; NUMBER OF SEQ ID NOS: 418
; SOFTWARE: Patentin version 3.3
; SEQ ID NO 248
; LENGTH: 3476
; TYPE: PRT
; ORGANISM: Schizosaccharomyces pombe
US-11-063-439-248

Query Match 2.2%; Score 12; DB 7; Length 3476;
Best Local Similarity 100.0%; Pred. No. 0.026;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 393 TTTT TTTT TTTT TTTT 404
DB 1779 TTTT TTTT TTTT TTTT 1790

RESULT 9

US-11-063-439-86
; Sequence 86, Application US/11063439
; Publication No. US20060147371A1
; GENERAL INFORMATION:
; APPLICANT: TUSZYNSKI, JACK A.
; APPLICANT: GREENWALD, HOWARD J.
; APPLICANT: CURRY, STEPHEN H.
; APPLICANT: GOSS, KENDRICK
; TITLE OF INVENTION: WATER-SOLUBLE COMPOUND
; FILE REFERENCE: 1034312-000027
; CURRENT APPLICATION NUMBER: US/11/063,439
; CURRENT FILING DATE: 2005-02-23
; PRIOR APPLICATION NUMBER: 10/878,905
; PRIOR FILING DATE: 2004-06-28
; PRIOR APPLICATION NUMBER: 10/923,615
; PRIOR FILING DATE: 2004-08-20

;; PRIOR APPLICATION NUMBER: 60/516,134
;; PRIOR FILING DATE: 2003-10-31
;; PRIOR APPLICATION NUMBER: 10/808,618
;; PRIOR FILING DATE: 2004-03-24
;; PRIOR APPLICATION NUMBER: 10/867,517
;; PRIOR FILING DATE: 2004-06-14
;; NUMBER OF SEQ ID NOS: 418
;; SOFTWARE: PatentIn version 3.3
;; SEQ ID NO: 86
;; LENGTH: 3481
;; TYPE: PRT
;; ORGANISM: Physarum polycephalum
US-11-063-439-86

Query Match 2.2%; Score 12; DB 7; Length 3481;
Best Local Similarity 100.0%; Pred. No. 0.026;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 393 TTTT TTTT TTTT TTTT 404
Db 1835 TTTT TTTT TTTT TTTT 1846

RESULT 10
US-11-063-439-10
;; Sequence 10, Application US/11063439
;; Publication No. US20060147371A1
;; GENERAL INFORMATION:

;; APPLICANT: TUSZYSKI, JACK A.
;; APPLICANT: GREENWALD, HOWARD J.
;; APPLICANT: CURRY, STEPHEN H.
;; APPLICANT: GOSS, KENDRICK

;; TITLE OF INVENTION: WATER-SOLUBLE COMPOUND

;; FILE REFERENCE: 1034312-000027

;; CURRENT APPLICATION NUMBER: US/11/063,439

;; CURRENT FILING DATE: 2005-02-23

;; PRIOR APPLICATION NUMBER: 10/878,905

;; PRIOR FILING DATE: 2004-06-28

;; PRIOR APPLICATION NUMBER: 10/923,615

;; PRIOR FILING DATE: 2004-08-20

;; PRIOR APPLICATION NUMBER: 60/516,134

;; PRIOR FILING DATE: 2003-10-31

;; PRIOR APPLICATION NUMBER: 10/808,618

;; PRIOR FILING DATE: 2004-03-24

;; PRIOR APPLICATION NUMBER: 10/867,517

;; NUMBER OF SEQ ID NOS: 418

;; SOFTWARE: PatentIn version 3.3

;; SEQ ID NO: 10

;; LENGTH: 3485

;; TYPE: PRT

;; ORGANISM: Hordeum vulgare

US-11-063-439-10

Query Match 2.2%; Score 12; DB 7; Length 3485;
Best Local Similarity 100.0%; Pred. No. 0.026;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 393 TTTT TTTT TTTT TTTT 404
Db 1834 TTTT TTTT TTTT TTTT 1845

RESULT 11
US-11-063-439-59

;; Sequence 59, Application US/11063439
;; Publication No. US20060147371A1
;; GENERAL INFORMATION:

;; APPLICANT: TUSZYSKI, JACK A.
;; APPLICANT: GREENWALD, HOWARD J.
;; APPLICANT: CURRY, STEPHEN H.
;; APPLICANT: GOSS, KENDRICK

;; TITLE OF INVENTION: WATER-SOLUBLE COMPOUND

US-11-063-439-59

;; FILE REFERENCE: 1034312-000027
;; CURRENT APPLICATION NUMBER: US/11/063,439
;; CURRENT FILING DATE: 2005-02-23
;; PRIOR APPLICATION NUMBER: 10/878,905
;; PRIOR FILING DATE: 2004-06-28
;; PRIOR APPLICATION NUMBER: 10/923,615
;; PRIOR FILING DATE: 2004-08-20
;; PRIOR APPLICATION NUMBER: 60/516,134
;; PRIOR FILING DATE: 2003-10-31
;; PRIOR APPLICATION NUMBER: 10/808,618
;; PRIOR FILING DATE: 2004-03-24
;; PRIOR APPLICATION NUMBER: 10/867,517
;; PRIOR FILING DATE: 2004-06-14
;; NUMBER OF SEQ ID NOS: 418
;; SOFTWARE: PatentIn version 3.3
;; SEQ ID NO: 59
;; LENGTH: 3485
;; TYPE: PRT
;; ORGANISM: Zea mays
US-11-063-439-59

Query Match 2.2%; Score 12; DB 7; Length 3485;
Best Local Similarity 100.0%; Pred. No. 0.026;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 393 TTTT TTTT TTTT TTTT 404
Db 1834 TTTT TTTT TTTT TTTT 1845

RESULT 12
US-11-063-439-76

;; Sequence 76, Application US/11063439
;; Publication No. US20060147371A1
;; GENERAL INFORMATION:

;; APPLICANT: TUSZYSKI, JACK A.
;; APPLICANT: GREENWALD, HOWARD J.
;; APPLICANT: CURRY, STEPHEN H.
;; APPLICANT: GOSS, KENDRICK

;; TITLE OF INVENTION: WATER-SOLUBLE COMPOUND

;; FILE REFERENCE: 1034312-000027

;; CURRENT APPLICATION NUMBER: US/11/063,439

;; CURRENT FILING DATE: 2005-02-23

;; PRIOR APPLICATION NUMBER: 10/878,905

;; PRIOR FILING DATE: 2004-06-28

;; PRIOR APPLICATION NUMBER: 10/923,615

;; PRIOR FILING DATE: 2004-08-20

;; PRIOR APPLICATION NUMBER: 60/516,134

;; PRIOR FILING DATE: 2003-10-31

;; PRIOR APPLICATION NUMBER: 10/808,618

;; PRIOR FILING DATE: 2004-03-24

;; PRIOR APPLICATION NUMBER: 10/867,517

;; NUMBER OF SEQ ID NOS: 418

;; SOFTWARE: PatentIn version 3.3

;; SEQ ID NO: 76

;; LENGTH: 3485

;; TYPE: PRT

;; ORGANISM: Physarum polycephalum

US-11-063-439-76

Query Match 2.2%; Score 12; DB 7; Length 3485;
Best Local Similarity 100.0%; Pred. No. 0.026;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 393 TTTT TTTT TTTT TTTT 404
Db 1838 TTTT TTTT TTTT TTTT 1849

RESULT 13
US-11-063-439-45

;; Sequence 45, Application US/11063439

Publication No. US20060147371A1
GENERAL INFORMATION:
APPLICANT: TUSZYNSKI, JACK A.
APPLICANT: GREENWALD, HOWARD J.
APPLICANT: CURRY, STEPHEN H.
APPLICANT: GOSS, KENDRICK
TITLE OF INVENTION: WATER-SOLUBLE COMPOUND
FILE REFERENCE: 1034312-000027
CURRENT APPLICATION NUMBER: US/11/063,439
CURRENT FILING DATE: 2005-02-23
PRIOR APPLICATION NUMBER: 10/878,905
PRIOR FILING DATE: 2004-06-28
PRIOR APPLICATION NUMBER: 10/923,615
PRIOR FILING DATE: 2004-08-20
PRIOR APPLICATION NUMBER: 60/516,134
PRIOR FILING DATE: 2003-10-31
PRIOR APPLICATION NUMBER: 10/808,618
PRIOR FILING DATE: 2004-03-24
PRIOR APPLICATION NUMBER: 10/867,517
PRIOR FILING DATE: 2004-06-14
NUMBER OF SEQ ID NOS: 418
SOFTWARE: PatentIn version 3.3
SEQ ID NO 45
LENGTH: 3486
TYPE: PRT
ORGANISM: Eleusine indica
US-11-063-439-45

Query Match 2.2%; Score 12; DB 7; Length 3486;
Best Local Similarity 100.0%; Pred. No. 0.026;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 393 TTTTTTTTTTS 404
Db 1834 TTTTTTTTTTS 1845

RESULT 14
US-11-063-439-56
Sequence 56, Application US/11063439
Publication No. US20060147371A1
GENERAL INFORMATION:
APPLICANT: TUSZYNSKI, JACK A.
APPLICANT: GREENWALD, HOWARD J.
APPLICANT: CURRY, STEPHEN H.
APPLICANT: GOSS, KENDRICK
TITLE OF INVENTION: WATER-SOLUBLE COMPOUND
FILE REFERENCE: 1034312-000027
CURRENT APPLICATION NUMBER: US/11/063,439
CURRENT FILING DATE: 2005-02-23
PRIOR APPLICATION NUMBER: 10/878,905
PRIOR FILING DATE: 2004-06-28
PRIOR APPLICATION NUMBER: 10/923,615
PRIOR FILING DATE: 2004-08-20
PRIOR APPLICATION NUMBER: 60/516,134
PRIOR FILING DATE: 2003-10-31
PRIOR APPLICATION NUMBER: 10/808,618
PRIOR FILING DATE: 2004-03-24
PRIOR APPLICATION NUMBER: 10/867,517
PRIOR FILING DATE: 2004-06-14
NUMBER OF SEQ ID NOS: 418
SOFTWARE: PatentIn version 3.3
SEQ ID NO 56
LENGTH: 3487
TYPE: PRT
ORGANISM: Zea mays
US-11-063-439-56

Query Match 2.2%; Score 12; DB 7; Length 3487;
Best Local Similarity 100.0%; Pred. No. 0.026;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 393 TTTTTTTTTTS 404

Db 1836 TTTTTTTTTTS 1847

RESULT 15
US-11-063-439-42
Sequence 42, Application US/11063439
Publication No. US20060147371A1
GENERAL INFORMATION:
APPLICANT: TUSZYNSKI, JACK A.
APPLICANT: GREENWALD, HOWARD J.
APPLICANT: CURRY, STEPHEN H.
APPLICANT: GOSS, KENDRICK
TITLE OF INVENTION: WATER-SOLUBLE COMPOUND
FILE REFERENCE: 1034312-000027
CURRENT APPLICATION NUMBER: US/11/063,439
CURRENT FILING DATE: 2005-02-23
PRIOR APPLICATION NUMBER: 10/878,905
PRIOR FILING DATE: 2004-06-28
PRIOR APPLICATION NUMBER: 10/923,615
PRIOR FILING DATE: 2004-08-20
PRIOR APPLICATION NUMBER: 60/516,134
PRIOR FILING DATE: 2003-10-31
PRIOR APPLICATION NUMBER: 10/808,618
PRIOR FILING DATE: 2004-03-24
PRIOR APPLICATION NUMBER: 10/867,517
PRIOR FILING DATE: 2004-06-14
NUMBER OF SEQ ID NOS: 418
SOFTWARE: PatentIn version 3.3
SEQ ID NO 42
LENGTH: 3488
TYPE: PRT
ORGANISM: Arabidopsis thaliana
US-11-063-439-42

Query Match 2.2%; Score 12; DB 7; Length 3488;
Best Local Similarity 100.0%; Pred. No. 0.026;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 393 TTTTTTTTTTS 404
Db 1836 TTTTTTTTTTS 1847

Search completed: August 1, 2006, 22:54:10
Job time : 34 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: August 1, 2006, 22:29:51 ; Search time 303 Seconds
(without alignments)
1703.494 Million cell updates/sec

Title: US-09-189-415D-11

Perfect score: 558
Sequence: 1 MFIGNLGNPNVNSIPAP.....SNSAVNTSNNPPAGSHRFV 558

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 2849598 seqs, 925015592 residues

Word size : 8

Total number of hits satisfying chosen parameters: 4155

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : UniProt_7.2.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	558	100.0	558	2	Q9R396_ECOLI
2	558	100.0	558	2	Q7DB77_ECOLI
3	340	60.9	558	2	O85506_ECOLI
4	223	40.0	558	2	Q4ZIM4_ECOLI
5	199	35.7	574	2	Q58188_ECOLI
6	28	5.0	547	2	Q7BHL5_GENTR
7	28	5.0	547	2	Q9ETI1_GENTR
8	28	5.0	547	2	Q9WXK1_ECOLI
9	22	3.9	538	2	O85508_ECOLI
10	22	3.9	538	2	Q47014_ECOLI
11	22	3.9	538	2	Q47016_ECOLI
12	22	3.9	538	2	Q58187_ECOLI
13	22	3.9	538	2	Q58189_ECOLI
14	22	3.9	538	2	Q58190_ECOLI
15	22	3.9	538	2	Q5KSP9_ECOLI
16	20	3.6	367	2	Q79C12_ECOLI
17	14	2.5	549	2	O50190_ECOLI
18	14	2.5	550	2	O52147_ECOLI
19	14	2.5	550	2	O5WMC9_ECOLI
20	14	2.5	551	2	O68258_ECOLI
21	14	2.5	551	2	Q4ZIM0_ECOLI
22	14	2.5	551	2	Q4ZIM1_ECOLI
23	14	2.5	552	2	Q9KMH9_ECOLI
24	13	2.3	163	2	O6GED0_STAR
25	13	2.3	165	2	O4CXK1_TRYCR
26	13	2.3	166	2	O2YYU1_STAB
27	13	2.3	166	2	O5HDQ5_STAC
28	13	2.3	166	2	O6G718_STAS
29	13	2.3	166	2	Q7A060_STAM
30	13	2.3	166	2	Q7A2K7_STAM
31	13	2.3	166	2	Q7A418_STAM

32	13	2.3	207	2	Q4DPK5_TRYCR	Q4dpk5_rypanosoma
33	13	2.3	271	2	O94491_DICDI	O94491_dicyosteli
34	13	2.3	498	2	O86J21_DICDI	O86J21_dicyosteli
35	13	2.3	514	2	O5SGG6_DICDI	O5SGG6_dicyosteli
36	13	2.3	517	2	O5ATR7_DICDI	O5ATR7_dicyosteli
37	13	2.3	665	2	O55G56_DICDI	O55G56_dicyosteli
38	13	2.3	951	2	O5SEY5_DICDI	O5SEY5_dicyosteli
39	13	2.3	1025	2	O4N714_THEPA	O4N714_theliera p
40	12	2.2	56	2	O8TIV8_DICDI	O8TIV8_dicyosteli
41	12	2.2	89	2	O9NTP9_TRYCR	O9NTP9_rypanosoma
42	12	2.2	106	2	O6WB03_TRYCR	O6WB03_rypanosoma
43	12	2.2	106	2	O6WB04_TRYCR	O6WB04_rypanosoma
44	12	2.2	106	2	O6WB05_TRYCR	O6WB05_rypanosoma
45	12	2.2	107	2	O6WB01_TRYCR	O6WB01_rypanosoma
46	12	2.2	107	2	O6WB08_TRYCR	O6WB08_rypanosoma
47	12	2.2	107	2	O9N1Q1_TRYCR	O9N1Q1_rypanosoma
48	12	2.2	108	2	O6WB02_TRYCR	O6WB02_rypanosoma
49	12	2.2	119	2	O4E362_TRYCR	O4E362_rypanosoma
50	12	2.2	119	2	O9N6G3_TRYCR	O9N6G3_rypanosoma
51	12	2.2	120	2	O9N1Q0_TRYCR	O9N1Q0_rypanosoma
52	12	2.2	121	2	O4E363_TRYCR	O4E363_rypanosoma
53	12	2.2	127	2	O4D956_TRYCR	O4D956_rypanosoma
54	12	2.2	130	2	O6WB06_TRYCR	O6WB06_rypanosoma
55	12	2.2	132	2	O9N1Q3_TRYCR	O9N1Q3_rypanosoma
56	12	2.2	133	2	O4D2Q6_TRYCR	O4D2Q6_rypanosoma
57	12	2.2	134	2	O4E364_TRYCR	O4E364_rypanosoma
58	12	2.2	139	2	O9CTQ5_ARATH	O9CTQ5_arabidopsis
59	12	2.2	141	2	O4CN03_TRYCR	O4CN03_rypanosoma
60	12	2.2	150	2	O8LEL8_ARATH	O8LEL8_arabidopsis
61	12	2.2	152	2	O4E366_TRYCR	O4E366_rypanosoma
62	12	2.2	188	2	O5CWP3_CRYHO	O5CWP3_cryptospori
63	12	2.2	192	2	O4D9V5_TRYCR	O4D9V5_rypanosoma
64	12	2.2	200	2	O9XX76_CABEL	O9XX76_caenorhabdi
65	12	2.2	205	2	O4CZJ6_TRYCR	O4CZJ6_rypanosoma
66	12	2.2	208	2	O4DVA1_TRYCR	O4DVA1_rypanosoma
67	12	2.2	208	2	O4E328_TRYCR	O4E328_rypanosoma
68	12	2.2	210	2	O4DE03_TRYCR	O4DE03_rypanosoma
69	12	2.2	210	2	O4DZB3_TRYCR	O4DZB3_rypanosoma
70	12	2.2	218	2	O4E365_TRYCR	O4E365_rypanosoma
71	12	2.2	229	2	O4DN80_TRYCR	O4DN80_rypanosoma
72	12	2.2	237	2	O54WY0_DICDI	O54WY0_dicyosteli
73	12	2.2	257	2	O54HH8_DICDI	O54HH8_dicyosteli
74	12	2.2	282	2	O5CEI0_CRYHO	O5CEI0_cryptospori
75	12	2.2	286	2	O54Y27_DICDI	O54Y27_dicyosteli
76	12	2.2	291	2	O94467_DICDI	O94467_dicyosteli
77	12	2.2	304	2	O54YH7_DICDI	O54YH7_dicyosteli
78	12	2.2	306	2	O4A3Y2_RUMFL	O4A3Y2_ruminooccu
79	12	2.2	333	2	O55DU4_DICDI	O55DU4_dicyosteli
80	12	2.2	336	1	RT09_CANAL	O94150_candida alb
81	12	2.2	336	2	O5A3F0_CANAL	O5A3F0_candida alb
82	12	2.2	337	2	O86K30_DICDI	O86K30_dicyosteli
83	12	2.2	340	2	O7P221_ANOGA	O7P221_anopheles g
84	12	2.2	344	2	O54ZU0_DICDI	O54ZU0_dicyosteli
85	12	2.2	359	2	O54M15_DICDI	O54M15_dicyosteli
86	12	2.2	364	2	O9W227_DROME	O9W227_drosophila
87	12	2.2	365	2	O55FX6_DICDI	O55FX6_dicyosteli
88	12	2.2	368	2	O54EAL_DICDI	O54EAL_dicyosteli
89	12	2.2	377	2	O5CF86_CRYHO	O5CF86_cryptospori
90	12	2.2	378	2	O55D73_DICDI	O55D73_dicyosteli
91	12	2.2	383	2	O55C29_DICDI	O55C29_dicyosteli
92	12	2.2	383	2	O5CSB3_CRYPV	O5CSB3_cryptospori
93	12	2.2	388	2	O61W24_CABER	O61W24_caenorhabdi
94	12	2.2	397	2	O54UG1_DICDI	O54UG1_dicyosteli
95	12	2.2	402	2	O59N14_CANAL	O59N14_candida alb
96	12	2.2	403	2	O44626_CABEL	O44626_caenorhabdi
97	12	2.2	402	2	O9XJ33_CVACA	O9XJ33_cyanidium c
98	12	2.2	432	2	O54RX5_DICDI	O54RX5_dicyosteli
99	12	2.2	438	2	O55FY4_DICDI	O55FY4_dicyosteli
100	12	2.2	439	2	O54Q08_DICDI	O54Q08_dicyosteli
101	12	2.2	444	2	O7S050_NEUCR	O7S050_neurospora
102	12	2.2	445	2	O8K3TE_MOUSE	O8K3TE_mus musculu
103	12	2.2	458	2	O9NA83_CABEL	O9NA83_caenorhabdi
104	12	2.2	465	2	O5CXW6_CRYPV	O5CXW6_cryptospori

105	12	2.2	467	2	05C191_CRYHO	05C191_cryptospori	178	12	2.2	860	2	054MY9_DICDI	054my9_dicystostei
106	12	2.2	484	2	02P0N7_GLOMR	02p0n7_glossina mo	179	12	2.2	863	2	054Q23_DICDI	054q23_dicystostei
107	12	2.2	486	2	07S0Z3_NEUCR	07s0z3_neurospora	180	12	2.2	865	2	054T59_DICDI	054t59_dicystostei
108	12	2.2	496	2	054S17_DICDI	054s17_dicystostei	181	12	2.2	872	2	02G257_9MYCE	02g257_dicystostei
109	12	2.2	500	2	054D28_DICDI	054d28_dicystostei	182	12	2.2	873	2	055GF7_DICDI	055gf7_dicystostei
110	12	2.2	517	2	054J27_DICDI	054j27_dicystostei	183	12	2.2	883	2	054K01_DICDI	054k01_dicystostei
111	12	2.2	518	2	054R71_DICDI	054r71_dicystostei	184	12	2.2	903	2	055GB5_DICDI	055gb5_dicystostei
112	12	2.2	518	2	054T16_DICDI	054t16_dicystostei	185	12	2.2	906	2	054I69_DICDI	054i69_dicystostei
113	12	2.2	519	2	054Q44_DICDI	054q44_dicystostei	186	12	2.2	908	2	054PK9_DICDI	054pk9_dicystostei
114	12	2.2	520	2	05CM96_CRYPV	05cm96_cryptospori	187	12	2.2	913	2	05CIN9_CRYHO	05cin9_cryptospori
115	12	2.2	524	1	SMOX2_SCHNA	05cm96_cryptospori	188	12	2.2	920	2	054T71_DICDI	054t71_dicystostei
116	12	2.2	540	2	07S0M6_DICDI	026601_schistosoma	189	12	2.2	923	2	054NA9_DICDI	054na9_dicystostei
117	12	2.2	542	2	054PF2_DICDI	07s0m6_dicystostei	190	12	2.2	925	2	05CSA0_CRYPV	05csa0_cryptospori
118	12	2.2	544	2	054U78_DICDI	054pf2_dicystostei	191	12	2.2	938	2	054F31_DICDI	054f31_dicystostei
119	12	2.2	559	2	083151_9BETA	054u78_dicystostei	192	12	2.2	941	2	054XY0_DICDI	054xy0_dicystostei
120	12	2.2	569	2	083183_9BETA	083151_murid herpe	193	12	2.2	953	2	054UG2_DICDI	054ug2_dicystostei
121	12	2.2	574	2	058SM3_DROME	083183_murid herpe	194	12	2.2	970	2	054M19_DICDI	054m19_dicystostei
122	12	2.2	576	2	0558Q9_DROME	058sm3_drosophila	195	12	2.2	978	2	054DF8_DICDI	054df8_dicystostei
123	12	2.2	582	2	054LR8_DICDI	0558q9_dicystostei	196	12	2.2	980	2	054T19_DICDI	054t19_dicystostei
124	12	2.2	588	2	055CB4_DICDI	054lr8_dicystostei	197	12	2.2	986	2	09N9W3_HAECCO	09n9w3_haemonchus
125	12	2.2	589	2	054UG6_DICDI	055cb4_dicystostei	198	12	2.2	994	2	0210B2_OSTOS	0210b2_ostertagia
126	12	2.2	604	2	04MN28_ASPFU	054ug6_dicystostei	199	12	2.2	994	2	0210B4_9BILA	0210b4_cooperia on
127	12	2.2	606	2	054H01_DICDI	04mn28_aspergillus	200	12	2.2	994	2	0210B5_9BILA	0210b5_cooperia on
128	12	2.2	616	2	054VU0_DICDI	054h01_dicystostei	201	12	2.2	1026	2	0210B3_OSTOS	0210b3_ostertagia
129	12	2.2	623	2	055EQ0_DICDI	054vu0_dicystostei	202	12	2.2	1026	1	GLT_DROME	0210b3_ostertagia
130	12	2.2	626	2	054X74_DICDI	055eq0_dicystostei	203	12	2.2	1032	2	055EM8_DICDI	055em8_drosophila
131	12	2.2	633	2	054IL6_DICDI	054x74_dicystostei	204	12	2.2	1034	2	055FL7_DICDI	055fl7_dicystostei
132	12	2.2	638	2	054F01_DICDI	054il6_dicystostei	205	12	2.2	1036	2	054HR5_DICDI	054hr5_dicystostei
133	12	2.2	639	2	054JL7_DICDI	054f01_dicystostei	206	12	2.2	1068	2	055DX7_DICDI	055dx7_dicystostei
134	12	2.2	641	2	054I00_DICDI	054j17_dicystostei	207	12	2.2	1076	2	054KES_DICDI	054kes_dicystostei
135	12	2.2	646	2	08MNK4_DICDI	054i00_dicystostei	208	12	2.2	1101	2	054Q10_DICDI	054q10_dicystostei
136	12	2.2	648	2	086AK1_DICDI	08mnk4_dicystostei	209	12	2.2	1110	2	054L20_DICDI	054l20_dicystostei
137	12	2.2	649	2	055B08_DICDI	086ak1_d similar t	210	12	2.2	1123	2	081811_DICDI	081811_dicystostei
138	12	2.2	649	2	08MUP9_DICDI	055b08_dicystostei	211	12	2.2	1125	2	054XQ2_DICDI	054xq2_dicystostei
139	12	2.2	656	2	054RH5_DICDI	08mup9_dicystostei	212	12	2.2	1125	2	05C1S2_CRYHO	05c1s2_cryptospori
140	12	2.2	659	2	094A11_ARA7H	054rh5_dicystostei	213	12	2.2	1176	2	054QM4_DICDI	054qm4_dicystostei
141	12	2.2	661	2	054I40_DICDI	094a11_arabidopsis	214	12	2.2	1185	2	0551B2_DICDI	0551b2_dicystostei
142	12	2.2	671	2	054KG1_DICDI	054i40_dicystostei	215	12	2.2	1186	2	054I94_DICDI	054i94_dicystostei
143	12	2.2	672	2	04R1P1_TETNG	054kg1_tetradon n	216	12	2.2	1199	2	05CEX1_CRYHO	05cex1_cryptospori
144	12	2.2	674	2	07YZH9_MONBE	04r1p1_monosiga dr	217	12	2.2	1202	2	05CFI7_CRYHO	05cfi7_cryptospori
145	12	2.2	693	2	054D77_DICDI	07yzh9_sulfolobus	218	12	2.2	1208	2	054SK8_DICDI	054sk8_dicystostei
146	12	2.2	697	2	096XH6_SULTO	054d77_dicystostei	219	12	2.2	1259	2	054YZ6_DICDI	054yz6_dicystostei
147	12	2.2	702	2	05A112_CANAL	096xh6_candida alb	220	12	2.2	1226	2	055GF2_DICDI	055gf2_dicystostei
148	12	2.2	706	2	09V615_DROME	05a112_drosophila	221	12	2.2	1247	2	09U9T1_DICDI	09u9t1_dicystostei
149	12	2.2	711	2	07YXX9_CRYPV	07yxx9_cryptospori	222	12	2.2	1247	2	09Y0T2_DICDI	09y0t2_dicystostei
150	12	2.2	717	2	08U1H5_PYRFU	08u1h5_pyrococcus	223	12	2.2	1256	2	0675X6_9UROC	0675x6_atriplexurea
151	12	2.2	718	2	09C6C5_ARA7H	09c6c5_arabidopsis	224	12	2.2	1258	2	086KB3_DICDI	086kb3_dicystostei
152	12	2.2	719	2	055GN6_DICDI	055gn6_dicystostei	225	12	2.2	1259	2	054P74_DICDI	054p74_dicystostei
153	12	2.2	719	2	05CH11_CRYHO	05ch11_cryptospori	226	12	2.2	1335	2	054FA3_DICDI	054fa3_dicystostei
154	12	2.2	722	2	09C700_ARA7H	05c700_arabidopsis	227	12	2.2	1268	2	08SSS5_DICDI	08sss5_dicystostei
155	12	2.2	725	2	08T124_DICDI	08t124_dicystostei	228	12	2.2	1297	2	08ETU8_DICDI	08etu8_dicystostei
156	12	2.2	735	2	086LA3_DICDI	086la3_dicystostei	229	12	2.2	1313	2	08ETU8_DICDI	08etu8_dicystostei
157	12	2.2	766	2	06Q7J0_DROME	06q7j0_drosophila	230	12	2.2	1326	2	08ETU9_DICDI	08eti9_dicystostei
158	12	2.2	769	2	054DG4_DICDI	054dg4_dicystostei	231	12	2.2	1328	2	0553P9_DICDI	0553p9_dicystostei
159	12	2.2	777	2	054UX1_DICDI	054ux1_dicystostei	232	12	2.2	1335	2	054F03_DICDI	054f03_dicystostei
160	12	2.2	780	2	055A63_DICDI	055a63_dicystostei	233	12	2.2	1339	2	054TU0_DICDI	054tu0_dicystostei
161	12	2.2	780	2	05CH10_CRYHO	05ch10_cryptospori	234	12	2.2	1347	2	054JL7_DICDI	054jl7_dicystostei
162	12	2.2	781	2	055OLA_DICDI	055ola_dicystostei	235	12	2.2	1365	2	055AD5_DICDI	055ad5_dicystostei
163	12	2.2	785	2	059P12_CANAL	059p12_candida alb	236	12	2.2	1385	2	054PY6_DICDI	054py6_dicystostei
164	12	2.2	798	2	054CX5_DICDI	054cx5_dicystostei	237	12	2.2	1387	2	055DN5_DICDI	055dn5_dicystostei
165	12	2.2	798	2	086K66_DICDI	086k66_dicystostei	238	12	2.2	1387	2	054QO1_DICDI	054qo1_dicystostei
166	12	2.2	802	1	XYND_RTMPL	053117_ruminococcu	239	12	2.2	1387	2	06H9K1_DICDI	06h9k1_dicystostei
167	12	2.2	802	2	09S310_RUNPL	09s310_ruminococcu	240	12	2.2	1419	2	054H40_DICDI	054h40_dicystostei
168	12	2.2	820	2	04CXG5_TRYCR	04cxg5_trypanosoma	241	12	2.2	1427	2	054UVZ_DICDI	054uvz_dicystostei
169	12	2.2	830	2	04DBZ0_TRYCR	04dbz0_trypanosoma	242	12	2.2	1438	2	086H10_DICDI	086h10_dicystostei
170	12	2.2	832	2	086H63_DICDI	086h63_dicystostei	243	12	2.2	1447	2	05CFG8_CRYHO	05cfg8_cryptospori
171	12	2.2	834	2	054Y73_DICDI	054y73_dicystostei	244	12	2.2	1459	2	054S42_DICDI	054s42_dicystostei
172	12	2.2	838	2	060WT0_CABBR	060wt0_caenorhabdi	245	12	2.2	1461	2	054KGB_DICDI	054kgb_dicystostei
173	12	2.2	843	2	0553B3_DICDI	0553b3_dicystostei	246	12	2.2	1461	2	086AZ9_DICDI	086az9_dicystostei
174	12	2.2	848	2	054XX5_DICDI	054xx5_dicystostei	247	12	2.2	1462	2	055OK3_DICDI	055ok3_dicystostei
175	12	2.2	855	2	0869V0_DICDI	0869v0_dicystostei	248	12	2.2	1497	2	054Q57_DICDI	054q57_dicystostei
176	12	2.2	855	2	054NA6_DICDI	054na6_dicystostei	249	12	2.2	1505	2	054VU4_DICDI	054vu4_dicystostei
177	12	2.2	860	2	023916_DICDI	023916_dicystostei	250	12	2.2	1506	2	054BKO_DICDI	054bko_dicystostei

397	11	2.0	146	2	Q6RUB8_SIVCZ	Q6rub8 chimpanzee	470	11	2.0	196	2	Q25947_PLAFA	Q25947 plasmidium
398	11	2.0	147	2	Q6RUE3_SIVCZ	Q6rue3 chimpanzee	471	11	2.0	196	2	Q25951_PLAFA	Q25951 plasmidium
399	11	2.0	147	2	Q6RUE8_SIVCZ	Q6rue8 chimpanzee	472	11	2.0	196	2	Q4E1B8_TRYCR	Q4E1B8 trypanosoma
400	11	2.0	148	2	Q61019_TRYCR	Q61019 trypanosoma	473	11	2.0	196	2	Q21C42_PLAFA	Q21C42 plasmidium
401	11	2.0	148	2	Q6WB00_TRYCR	Q6wb00 trypanosoma	474	11	2.0	197	2	Q35409_PLAFA	Q35409 plasmidium
402	11	2.0	148	2	Q6RUC1_SIVCZ	Q6ruc1 chimpanzee	475	11	2.0	197	2	Q4E5Y0_TRYCR	Q4E5Y0 trypanosoma
403	11	2.0	150	2	Q4CTF6_TRYCR	Q4ctf6 trypanosoma	476	11	2.0	198	2	Q35471_PLAFA	Q35471 plasmidium
404	11	2.0	150	2	Q9BJ06_PLAFA	Q9bj06 plasmidium	477	11	2.0	198	2	Q4DCY9_TRYCR	Q4dcy9 trypanosoma
405	11	2.0	150	2	Q9BJ07_PLAFA	Q9bj07 plasmidium	478	11	2.0	198	2	Q4DCC1_TRYCR	Q4dcc1 trypanosoma
406	11	2.0	150	2	Q9GOX3_PLAFA	Q9gqx3 plasmidium	479	11	2.0	199	2	Q4DH81_TRYCR	Q4dh81 trypanosoma
407	11	2.0	150	2	Q9VTF1_DROME	Q9vtf1 drosophila	480	11	2.0	200	2	Q4D7M9_TRYCR	Q4d7m9 trypanosoma
408	11	2.0	152	2	Q5C045_SCHJA	Q5c045 schistosoma	481	11	2.0	200	2	Q4DE08_TRYCR	Q4de08 trypanosoma
409	11	2.0	155	2	Q4CUV4_TRYCR	Q4cu44 trypanosoma	482	11	2.0	201	2	Q4DUY8_TRYCR	Q4duy8 trypanosoma
410	11	2.0	155	2	Q6USF5_PLAFA	Q6usf5 plasmidium	483	11	2.0	201	2	Q4E295_TRYCR	Q4e295 trypanosoma
411	11	2.0	157	2	Q55AA4_DICDI	Q55aa4 dictyosteli	484	11	2.0	201	2	Q4E3V4_TRYCR	Q4e3v4 trypanosoma
412	11	2.0	160	2	Q94669_PLAFA	Q94669 plasmidium	485	11	2.0	202	2	Q01615_PWECA	Q01615 pneumocysti
413	11	2.0	162	2	Q4D2K9_TRYCR	Q4d2k9 trypanosoma	486	11	2.0	202	2	Q25952_PLAFA	Q25952 plasmidium
414	11	2.0	164	2	Q9BJ05_PLAFA	Q9bj05 plasmidium	487	11	2.0	202	2	Q4D1F2_TRYCR	Q4d1f2 trypanosoma
415	11	2.0	165	2	Q94671_PLAFA	Q94671 plasmidium	488	11	2.0	202	2	Q4DLD7_TRYCR	Q4dlld7 trypanosoma
416	11	2.0	172	2	Q9VZB8_DROME	Q9vzb8 drosophila	489	11	2.0	202	2	Q4E044_TRYCR	Q4e044 trypanosoma
417	11	2.0	173	2	Q4CSC0_TRYCR	Q4csc0 trypanosoma	490	11	2.0	203	2	Q15806_PLAFA	Q15806 plasmidium
418	11	2.0	174	2	Q4CPUS_TRYCR	Q4cpus trypanosoma	491	11	2.0	203	2	Q353U5_PLAFA	Q353u5 plasmidium
419	11	2.0	174	2	Q4CX71_TRYCR	Q4cx71 trypanosoma	492	11	2.0	203	2	Q4D6C3_TRYCR	Q4d6c3 trypanosoma
420	11	2.0	174	2	Q4E216_TRYCR	Q4e216 trypanosoma	493	11	2.0	203	2	Q9U0C0_PLAFA	Q9u0c0 plasmidium
421	11	2.0	175	2	Q4E0E7_TRYCR	Q4e0e7 trypanosoma	494	11	2.0	203	2	Q9W2W1_DROME	Q9w2w1 drosophila
422	11	2.0	175	2	Q5TWM1_ANOGA	Q5twm1 anophelis	495	11	2.0	204	2	Q91VA3_ARATH	Q91va3 arabidopsis
423	11	2.0	177	2	Q25713_PLAFA	Q25713 plasmidium	496	11	2.0	204	2	Q4E454_TRYCR	Q4e454 arabidopsis
424	11	2.0	178	2	Q4DWM8_TRYCR	Q4dwm8 trypanosoma	497	11	2.0	204	2	Q8H7E2_ARATH	Q8h7e2 arabidopsis
425	11	2.0	179	2	Q9X240_PLAFA	Q9x240 plasmidium	498	11	2.0	205	2	Q15777_TRYCR	Q15777 trypanosoma
426	11	2.0	180	2	Q4DXN3_TRYCR	Q4dxn3 trypanosoma	499	11	2.0	205	2	Q15911_DICDI	Q15911 dictyosteli
427	11	2.0	180	2	Q4E284_TRYCR	Q4e284 trypanosoma	500	11	2.0	205	2	Q35405_PLAFA	Q35405 plasmidium
428	11	2.0	181	2	Q4D0G2_TRYCR	Q4d0g2 trypanosoma	501	11	2.0	205	2	Q4D8A4_TRYCR	Q4d8a4 trypanosoma
429	11	2.0	182	2	Q4D178_TRYCR	Q4d178 trypanosoma	502	11	2.0	205	2	Q4DHM9_TRYCR	Q4dhm9 trypanosoma
430	11	2.0	182	2	Q4DNF5_TRYCR	Q4dnps trypanosoma	503	11	2.0	205	2	Q4DYF0_TRYCR	Q4dyf0 trypanosoma
431	11	2.0	182	2	Q4TYV8_PLAFA	Q4tyv8 plasmidium	504	11	2.0	205	2	Q4DYF7_TRYCR	Q4dyf7 trypanosoma
432	11	2.0	183	1	AAC1_DICDI	P14195 dictyosteli	505	11	2.0	205	2	Q4DZB5_TRYCR	Q4dzbs trypanosoma
433	11	2.0	183	2	Q4DXZ9_TRYCR	Q4dxz9 trypanosoma	506	11	2.0	205	2	Q4E0E2_TRYCR	Q4e0e2 trypanosoma
434	11	2.0	183	2	Q4E068_TRYCR	Q4e068 trypanosoma	507	11	2.0	206	2	Q4DJK8_TRYCR	Q4djkh8 trypanosoma
435	11	2.0	183	2	Q6UA97_PLAFA	Q6ua97 plasmidium	508	11	2.0	206	2	Q9WPF9_GNUCL	Q9wpf9 ameacta alb
436	11	2.0	183	2	Q9U0B5_PLAFA	Q9u0b5 plasmidium	509	11	2.0	207	2	Q25701_PLAFA	Q25701 plasmidium
437	11	2.0	183	2	Q9U0B6_PLAFA	Q9u0b6 plasmidium	510	11	2.0	207	2	Q35313_PLAFA	Q35313 plasmidium
438	11	2.0	184	2	Q4E495_TRYCR	Q4e495 trypanosoma	511	11	2.0	207	2	Q4DGC3_TRYCR	Q4dgc3 trypanosoma
439	11	2.0	184	2	Q9TY34_PLAFA	Q9ty34 plasmidium	512	11	2.0	208	2	Q25949_PLAFA	Q25949 plasmidium
440	11	2.0	184	2	Q9U0B8_PLAFA	Q9u0b8 plasmidium	513	11	2.0	208	2	Q4DF66_TRYCR	Q4dff6 trypanosoma
441	11	2.0	184	2	Q9U0B9_PLAFA	Q9u0b9 plasmidium	514	11	2.0	208	2	Q4E2A8_TRYCR	Q4e2a8 trypanosoma
442	11	2.0	185	2	Q354V4_PLAFA	Q3e4v4 plasmidium	515	11	2.0	209	2	Q61055_TRYCR	Q61055 trypanosoma
443	11	2.0	187	2	Q5CYR7_CRPV	Q5cyr7 cryptospori	516	11	2.0	209	2	Q4DYW8_TRYCR	Q4dyw8 trypanosoma
444	11	2.0	189	2	Q4D1J5_TRYCR	Q4d1j5 trypanosoma	517	11	2.0	209	2	Q4DYV5_TRYCR	Q4dyv5 trypanosoma
445	11	2.0	189	2	Q4D373_TRYCR	Q4d373 trypanosoma	518	11	2.0	210	2	Q4D2X3_TRYCR	Q4d2x3 trypanosoma
446	11	2.0	189	2	Q4DUV1_TRYCR	Q4duv1 trypanosoma	519	11	2.0	210	2	Q4E1A1_TRYCR	Q4e1a1 trypanosoma
447	11	2.0	190	2	Q35605_PLAFA	Q35605 plasmidium	520	11	2.0	210	2	Q4E2H0_TRYCR	Q4e2h0 trypanosoma
448	11	2.0	190	2	Q4E2F3_TRYCR	Q4e2f3 trypanosoma	521	11	2.0	210	2	Q9Y025_TRYCR	Q9y025 trypanosoma
449	11	2.0	190	2	Q4E3V9_TRYCR	Q4e3v9 trypanosoma	522	11	2.0	211	2	Q354U6_PLAFA	Q354u6 plasmidium
450	11	2.0	190	2	Q54VT7_DICDI	Q54vt7 dictyosteli	523	11	2.0	211	2	Q4D2P0_TRYCR	Q4d2p0 trypanosoma
451	11	2.0	191	2	Q354V3_PLAFA	Q354v3 plasmidium	524	11	2.0	211	2	Q4D2P2_TRYCR	Q4d2p2 trypanosoma
452	11	2.0	191	2	Q4DU09_TRYCR	Q4du09 trypanosoma	525	11	2.0	211	2	Q4E3U1_TRYCR	Q4e3u1 trypanosoma
453	11	2.0	191	2	Q4E1V3_TRYCR	Q4e1v3 trypanosoma	526	11	2.0	211	2	Q4E452_TRYCR	Q4e452 trypanosoma
454	11	2.0	191	2	Q4E3V3_TRYCR	Q4e3v3 trypanosoma	527	11	2.0	212	2	Q4DTB8_TRYCR	Q4dtb8 trypanosoma
455	11	2.0	191	2	Q54WLS_DICDI	Q54wls dictyosteli	528	11	2.0	213	2	Q4E5T4_TRYCR	Q4e5t4 trypanosoma
456	11	2.0	191	2	Q7OEP3_ANOGA	Q7oep3 anophelis	529	11	2.0	214	2	Q4DNG2_TRYCR	Q4dng2 trypanosoma
457	11	2.0	191	2	Q9U0C1_PLAFA	Q9u0c1 plasmidium	530	11	2.0	215	2	Q4CUP6_TRYCR	Q4culf6 trypanosoma
458	11	2.0	192	2	Q4DVX1_TRYCR	Q4dvx1 trypanosoma	531	11	2.0	215	2	Q4DNP4_TRYCR	Q4dnp4 trypanosoma
459	11	2.0	192	2	Q5DTE9_CAEEL	Q5dtes caenorhabdi	532	11	2.0	215	2	Q4DTH6_TRYCR	Q4dth6 trypanosoma
460	11	2.0	193	2	Q4DTJ2_TRYCR	Q4dtj2 trypanosoma	533	11	2.0	216	2	Q4CUB8_TRYCR	Q4cub6 trypanosoma
461	11	2.0	193	2	Q4E1H8_TRYCR	Q4e1h8 trypanosoma	534	11	2.0	216	2	Q4E3U4_TRYCR	Q4e3u4 trypanosoma
462	11	2.0	193	2	Q4E1J1_TRYCR	Q4e1j1 trypanosoma	535	11	2.0	216	2	Q962W6_TRYCR	Q962w6 trypanosoma
463	11	2.0	193	2	Q4E1Z7_TRYCR	Q4e1z7 trypanosoma	536	11	2.0	217	1	SGS3_DROSI	P13729 drosophila
464	11	2.0	193	2	Q4E230_TRYCR	Q4e230 trypanosoma	537	11	2.0	217	2	Q4E2A5_TRYCR	Q4e2a5 trypanosoma
465	11	2.0	194	2	Q4DVT2_TRYCR	Q4dvt2 trypanosoma	538	11	2.0	219	2	Q4COP1_TRYCR	Q4cgef1 trypanosoma
466	11	2.0	194	2	Q4DMD7_TRYCR	Q4dmd7 trypanosoma	539	11	2.0	219	2	Q4DC47_TRYCR	Q4dc47 trypanosoma
467	11	2.0	194	2	Q4E1Z2_TRYCR	Q4e1z2 trypanosoma	540	11	2.0	220	2	Q4DXA5_TRYCR	Q4dxa5 trypanosoma
468	11	2.0	194	2	Q54X00_DICDI	Q54x00 dictyosteli	541	11	2.0	220	2	Q9U0B3_PLAFA	Q9u0b3 plasmidium
469	11	2.0	195	2	Q93948_CANAL	Q93948 candida alb	542	11	2.0	222	2	Q45U78_CAELE	Q45u78 caenorhabdi

543	1.1	2.0	222	2	Q4DB53_TRYCR	Q4db53	trypanosoma	616	1.1	2.0	252	2	Q4G1J2_EPRST	Q4g1j2	epitactretus
544	1.1	2.0	222	2	Q4E4W9_TRYCR	Q4e4w9	trypanosoma	617	1.1	2.0	252	2	Q4G1J7_EPRST	Q4g1j7	epitactretus
545	1.1	2.0	223	2	Q4CZ01_TRYCR	Q4cz01	trypanosoma	618	1.1	2.0	257	2	Q5A4D2_CANAL	Q5a4d2	candida alb
546	1.1	2.0	223	2	Q4E441_TRYCR	Q4e441	trypanosoma	619	1.1	2.0	258	2	Q4G1I0_EPRST	Q4g1i0	epitactretus
547	1.1	2.0	224	2	Q4E045_TRYCR	Q4e045	trypanosoma	620	1.1	2.0	258	2	Q4G1J0_EPRST	Q4g1j0	epitactretus
548	1.1	2.0	224	2	Q4E3D1_TRYCR	Q4e3d1	trypanosoma	621	1.1	2.0	258	2	Q4G1J6_EPRST	Q4g1j6	epitactretus
549	1.1	2.0	224	2	Q4E428_TRYCR	Q4e428	trypanosoma	622	1.1	2.0	259	2	Q2YE09_EPRST	Q2ye09	epitactretus
550	1.1	2.0	225	2	Q4D496_TRYCR	Q4d496	trypanosoma	623	1.1	2.0	259	2	Q2YE12_EPRST	Q2ye12	epitactretus
551	1.1	2.0	225	2	Q4DAF4_TRYCR	Q4daf4	trypanosoma	624	1.1	2.0	259	2	Q2YE22_EPRST	Q2ye22	epitactretus
552	1.1	2.0	226	2	Q4DA16_TRYCR	Q4da16	trypanosoma	625	1.1	2.0	259	2	Q2YE23_EPRST	Q2ye23	epitactretus
553	1.1	2.0	227	2	Q4DDU6_TRYCR	Q4ddu6	trypanosoma	626	1.1	2.0	259	2	Q4G1H7_EPRST	Q4g1h7	epitactretus
554	1.1	2.0	227	2	Q4DF34_TRYCR	Q4df34	trypanosoma	627	1.1	2.0	259	2	Q4G1J3_EPRST	Q4g1j3	epitactretus
555	1.1	2.0	227	2	Q5BWT7_SCHJA	Q5bwt7	schistosoma	628	1.1	2.0	259	2	Q4G1J4_EPRST	Q4g1j4	epitactretus
556	1.1	2.0	228	2	Q4D389_TRYCR	Q4d389	trypanosoma	629	1.1	2.0	259	2	Q4G1J8_EPRST	Q4g1j8	epitactretus
557	1.1	2.0	228	2	Q4DDU9_TRYCR	Q4ddu9	trypanosoma	630	1.1	2.0	259	2	Q4G1K1_EPRST	Q4g1k1	epitactretus
558	1.1	2.0	228	2	Q4DUE0_TRYCR	Q4due0	trypanosoma	631	1.1	2.0	259	2	Q4G1M4_DICDI	Q4g1m4	epitactretus
559	1.1	2.0	229	2	Q4DBA1_TRYCR	Q4dba1	trypanosoma	632	1.1	2.0	259	2	Q4G1H5_EPRST	Q4g1h5	epitactretus
560	1.1	2.0	229	2	Q9V1A7_DROME	Q9v1a7	dirosophila	633	1.1	2.0	260	2	Q2YE05_EPRST	Q2ye05	epitactretus
561	1.1	2.0	230	2	Q4CZ78_TRYCR	Q4cz78	trypanosoma	634	1.1	2.0	260	2	Q4G1K7_EPRST	Q4g1k7	epitactretus
562	1.1	2.0	230	2	Q4DD08_TRYCR	Q4dd08	trypanosoma	635	1.1	2.0	260	2	Q81T83_PLAFA	Q81t83	plasmodium
563	1.1	2.0	230	2	Q4D9A8_TRYCR	Q4d9a8	trypanosoma	636	1.1	2.0	261	2	Q4G1J1_EPRST	Q4g1j1	epitactretus
564	1.1	2.0	230	2	Q4D065_TRYCR	Q4d065	trypanosoma	637	1.1	2.0	262	2	Q2YE11_EPRST	Q2ye11	epitactretus
565	1.1	2.0	230	2	Q4E3U7_TRYCR	Q4e3u7	trypanosoma	638	1.1	2.0	262	2	Q2YE18_EPRST	Q2ye18	epitactretus
566	1.1	2.0	231	2	Q4CZM6_TRYCR	Q4czm6	trypanosoma	639	1.1	2.0	262	2	Q4G1K0_EPRST	Q4g1k0	epitactretus
567	1.1	2.0	231	2	Q4DLI6_TRYCR	Q4dli6	trypanosoma	640	1.1	2.0	264	2	Q9E920_PLAFA	Q9e920	plasmodium
568	1.1	2.0	231	2	Q4E046_TRYCR	Q4e046	trypanosoma	641	1.1	2.0	272	2	Q5SGS7_DICDI	Q5sgs7	dicystostelli
569	1.1	2.0	231	2	Q4E074_TRYCR	Q4e074	trypanosoma	642	1.1	2.0	274	1	MSA2_PLAFA	MSA2	plasma
570	1.1	2.0	231	2	Q5BWA9_SCHJA	Q5bwa9	schistosoma	643	1.1	2.0	274	1	Q7SC44_NEUCR	Q7sc44	neutrospora
571	1.1	2.0	232	2	Q4CP74_TRYCR	Q4cp74	trypanosoma	644	1.1	2.0	275	2	Q4G1J5_EPRST	Q4g1j5	epitactretus
572	1.1	2.0	232	2	Q4CRO1_TRYCR	Q4cro1	trypanosoma	645	1.1	2.0	276	2	Q2YE25_EPRST	Q2ye25	epitactretus
573	1.1	2.0	232	2	Q4CUI2_TRYCR	Q4cui2	trypanosoma	646	1.1	2.0	277	2	Q8E1C7_DICDI	Q8e1c7	dicystostelli
574	1.1	2.0	232	2	Q4CYS1_TRYCR	Q4cys1	trypanosoma	647	1.1	2.0	278	2	Q2S862_PLAFA	Q2s862	plasmodium
575	1.1	2.0	232	2	Q4D0A6_TRYCR	Q4d0a6	trypanosoma	648	1.1	2.0	278	2	Q4H3L9_CIOIN	Q4h3l9	ciona intes
576	1.1	2.0	232	2	Q4DQ51_TRYCR	Q4dq51	trypanosoma	649	1.1	2.0	280	2	Q17639_CABEL	Q17639	caenorhabdi
577	1.1	2.0	232	2	Q4DYU8_TRYCR	Q4dyu8	trypanosoma	650	1.1	2.0	281	2	Q5AND8_DICDI	Q5and8	dicystostelli
578	1.1	2.0	233	2	Q4CVU7_TRYCR	Q4cvu7	trypanosoma	651	1.1	2.0	282	2	Q4G1I6_EPRST	Q4g1i6	epitactretus
579	1.1	2.0	234	2	Q4G1H3_EPRST	Q4g1h3	epitactretus	652	1.1	2.0	283	2	Q2YE14_EPRST	Q2ye14	epitactretus
580	1.1	2.0	234	2	Q4G1K3_EPRST	Q4g1k3	epitactretus	653	1.1	2.0	283	2	Q2YE20_EPRST	Q2ye20	epitactretus
581	1.1	2.0	235	2	Q4B5I9_TRYCR	Q4b5i9	trypanosoma	654	1.1	2.0	283	2	Q4G1H4_EPRST	Q4g1h4	epitactretus
582	1.1	2.0	235	2	Q4B5J0_TRYCR	Q4b5j0	trypanosoma	655	1.1	2.0	283	2	Q4G1H8_EPRST	Q4g1h8	epitactretus
583	1.1	2.0	235	2	Q4DBA2_TRYCR	Q4dba2	trypanosoma	656	1.1	2.0	283	2	Q4G1I1_EPRST	Q4g1i1	epitactretus
584	1.1	2.0	235	2	Q4G1H9_EPRST	Q4g1h9	epitactretus	657	1.1	2.0	283	2	Q61P66_CABRR	Q61p66	caenorhabdi
585	1.1	2.0	235	2	Q4G1I5_EPRST	Q4g1i5	epitactretus	658	1.1	2.0	283	2	Q8E1L5_DICDI	Q8e1l5	dicystostelli
586	1.1	2.0	235	2	Q7SDS9_NEUCR	Q7sds9	neutrospora	659	1.1	2.0	284	2	Q2O202_CABEL	Q2o202	caenorhabdi
587	1.1	2.0	236	2	Q5CSJ9_CRIPIV	Q5csj9	cryptospori	660	1.1	2.0	284	2	Q2YE15_EPRST	Q2ye15	epitactretus
588	1.1	2.0	237	2	Q2YE13_EPRST	Q2ye13	epitactretus	661	1.1	2.0	286	2	Q2YE07_EPRST	Q2ye07	epitactretus
589	1.1	2.0	237	2	Q4CUN2_TRYCR	Q4cun2	trypanosoma	662	1.1	2.0	286	2	Q2YE17_EPRST	Q2ye17	epitactretus
590	1.1	2.0	237	2	Q4CUT4_TRYCR	Q4cut4	trypanosoma	663	1.1	2.0	287	2	Q5AHP7_DICDI	Q5ahp7	dicystostelli
591	1.1	2.0	237	2	Q4DBN4_TRYCR	Q4dbn4	trypanosoma	664	1.1	2.0	288	2	Q4DNP7_TRYCR	Q4dnp7	trypanosoma
592	1.1	2.0	237	2	Q4DYV6_TRYCR	Q4dyv6	trypanosoma	665	1.1	2.0	291	2	Q54GV9_DICDI	Q54gv9	dicystostelli
593	1.1	2.0	238	2	Q4DBE4_TRYCR	Q4dbe4	trypanosoma	666	1.1	2.0	293	2	Q2S785_PLAFA	Q2s785	plasmodium
594	1.1	2.0	239	2	Q9WZRS_DROME	Q9wzrs	dirosophila	667	1.1	2.0	293	2	Q54P39_DICDI	Q54p39	dicystostelli
595	1.1	2.0	242	2	Q4CNE3_TRYCR	Q4cne3	trypanosoma	668	1.1	2.0	293	2	Q556R4_DICDI	Q556r4	dicystostelli
596	1.1	2.0	242	2	Q4CVN9_TRYCR	Q4cvn9	trypanosoma	669	1.1	2.0	295	2	Q9X2H6_MOUSE	Q9x2h6	mus musculu
597	1.1	2.0	242	2	Q4DVB8_TRYCR	Q4dwb8	trypanosoma	670	1.1	2.0	295	2	Q9QYI6_MOUSE	Q9qyi6	mus musculu
598	1.1	2.0	244	2	Q5ABR0_DICDI	Q5abr0	dicystostelli	671	1.1	2.0	296	2	Q4G1H6_EPRST	Q4g1h6	epitactretus
599	1.1	2.0	244	2	Q9U206_CABEL	Q9u206	caenorhabdi	672	1.1	2.0	297	2	Q61VC3_CABRR	Q61vc3	caenorhabdi
600	1.1	2.0	245	2	Q564Z3_CABEL	Q564z3	caenorhabdi	673	1.1	2.0	299	2	Q5A3W8_CANAL	Q5a3w8	candida alb
601	1.1	2.0	245	2	Q5BZW6_SCHJA	Q5bzw6	schistosoma	674	1.1	2.0	299	2	Q2YD23_EPRST	Q2yd23	epitactretus
602	1.1	2.0	245	2	Q9XWP2_CABEL	Q9xwp2	caenorhabdi	675	1.1	2.0	299	2	Q4DUV4_TRYCR	Q4duv4	trypanosoma
603	1.1	2.0	246	2	Q76P06_DICDI	Q76p06	dicystostelli	676	1.1	2.0	302	2	Q4G1I9_EPRST	Q4g1i9	epitactretus
604	1.1	2.0	247	2	Q54GM3_DICDI	Q54gm3	dicystostelli	677	1.1	2.0	302	2	Q9S023_DROME	Q9s023	dirosophila
605	1.1	2.0	248	2	Q2P098_CABEL	Q2p098	caenorhabdi	678	1.1	2.0	303	2	Q4QGL0_LEIIMA	Q4qgl0	leiishmania
606	1.1	2.0	248	2	Q2YE16_EPRST	Q2ye16	epitactretus	679	1.1	2.0	303	2	Q91IM9_WSCIV	Q91im9	white spot
607	1.1	2.0	248	2	Q4G1K4_EPRST	Q4g1k4	epitactretus	680	1.1	2.0	304	1	YO0B_CABEL	YO0B	CABEL
608	1.1	2.0	248	2	Q54R19_DICDI	Q54r19	dicystostelli	681	1.1	2.0	306	2	Q2YDZ6_EPRST	Q2ydz6	epitactretus
609	1.1	2.0	251	2	Q2YEO3_EPRST	Q2ye03	epitactretus	682	1.1	2.0	306	2	Q2YDZ7_EPRST	Q2ydz7	epitactretus
610	1.1	2.0	251	2	Q2YB24_EPRST	Q2yb24	epitactretus	683	1.1	2.0	306	2	Q2YE10_EPRST	Q2ye10	epitactretus
611	1.1	2.0	251	2	Q4G1I7_EPRST	Q4g1i7	epitactretus	684	1.1	2.0	306	2	Q2YE21_EPRST	Q2ye21	epitactretus
612	1.1	2.0	252	2	Q4G1K2_EPRST	Q4g1k2	epitactretus	685	1.1	2.0	306	2	Q4G1K6_EPRST	Q4g1k6	epitactretus
613	1.1	2.0	252	2	Q4G1I2_EPRST	Q4g1i2	epitactretus	686	1.1	2.0	306	2	Q6Z4Q3_CABRR	Q6z4q3	caenorhabdi
614	1.1	2.0	252	2	Q4G1I4_EPRST	Q4g1i4	epitactretus	687	1.1	2.0	306	2	Q9QYI4_MOUSE	Q9qyi4	mus musculu
615	1.1	2.0	252	2	Q4G1I8_EPRST	Q4g1i8	epitactretus	688	1.1	2.0	307	2	Q2YEO0_EPRST	Q2ye00	epitactretus

689	11	2.0	307	2	Q2YE04_EPTST	Q2YE04_eptretus	762	11	2.0	399	2	Q66UX1_DICD1	Q66jx1_dictyosteli
690	11	2.0	307	2	Q2YE19_EPTST	Q2YE19_eptretus	763	11	2.0	400	2	Q5AL15_CANAL	Q5al15_candida alb
691	11	2.0	308	2	Q2YD23_EPTST	Q2Yd23_eptretus	764	11	2.0	401	2	Q54UJ3_DICD1	Q54uj3_dictyosteli
692	11	2.0	308	2	Q2YD25_EPTST	Q2Yd25_eptretus	765	11	2.0	402	2	Q54JN3_DICD1	Q54jn3_dictyosteli
693	11	2.0	308	2	Q4G113_EPTST	Q4G113_eptretus	766	11	2.0	403	2	Q61GN2_CAEBR	Q61gn2_caenothabdi
694	11	2.0	309	2	Q2YD28_EPTST	Q2Yd28_eptretus	767	11	2.0	404	2	Q541J3_DICD1	Q541j3_dictyosteli
695	11	2.0	312	2	Q01824_PNECA	Q01824_pneumocysti	768	11	2.0	404	2	Q66SB6_DROVI	Q66sb6_drosophila
696	11	2.0	312	2	Q6VZ55_CNPV	Q6vz55_canarypox v	769	11	2.0	408	2	Q4WVR0_ASFPV	Q4wvr0_aspergillus
697	11	2.0	316	2	Q05BJ7_DICD1	Q05bj7_dictyosteli	770	11	2.0	410	2	Q5CRW6_CRYPV	Q5crw6_cryptospori
698	11	2.0	320	2	Q0U448_DICD1	Q0u448_dictyosteli	771	11	2.0	410	2	Q7Q956_ANOGA	Q7q956_anopheles g
699	11	2.0	322	2	Q0SXG9_DROME	Q0sxg9_drosophila	772	11	2.0	417	2	Q01760_PNECA	Q01760_pneumocysti
700	11	2.0	323	2	Q2YEO8_EPTST	Q2Yeo8_eptretus	773	11	2.0	417	2	Q54161_DICD1	Q54161_dictyosteli
701	11	2.0	323	2	Q2YEO8_EPTST	Q2Yeo8_eptretus	774	11	2.0	418	2	Q54SS8_DICD1	Q54ss8_dictyosteli
702	11	2.0	323	2	Q4E2C2_TRYCR	Q4E2c2_rypanosoma	775	11	2.0	420	2	Q5CFT4_CRYHO	Q5cft4_cryptospori
703	11	2.0	323	2	Q06AP2_DICD1	Q06ap2_dictyosteli	776	11	2.0	422	2	Q5ALN9_CANAL	Q5aln9_candida alb
704	11	2.0	327	2	Q25134_LEIMA	Q25134_leishmania	777	11	2.0	422	2	Q91BH8_NPVST	Q91bh8_spodoptera
705	11	2.0	327	2	Q54UX2_DICD1	Q54ux2_dictyosteli	778	11	2.0	423	2	Q55DC5_DICD1	Q55dc5_dictyosteli
706	11	2.0	330	2	Q2YD29_EPTST	Q2Yd29_eptretus	779	11	2.0	423	2	Q55EC6_DICD1	Q55ec6_dictyosteli
707	11	2.0	331	2	Q2YD24_EPTST	Q2Yd24_eptretus	780	11	2.0	425	2	Q15755_DICD1	Q15755_dictyosteli
708	11	2.0	331	2	Q55E25_DICD1	Q55e25_dictyosteli	781	11	2.0	425	2	Q54ZM4_DICD1	Q54zm4_dictyosteli
709	11	2.0	331	2	Q55FM4_DICD1	Q55fm4_dictyosteli	782	11	2.0	425	2	Q55BA9_DICD1	Q55ba9_dictyosteli
710	11	2.0	331	2	Q7PSS1_ANOGA	Q7pss1_anopheles g	783	11	2.0	425	2	Q06AP8_DICD1	Q06ap8_dictyosteli
711	11	2.0	334	2	Q55927_DICD1	Q55927_dictyosteli	784	11	2.0	425	2	Q9CZ25_CAEBL	Q9cz25_caenothabdi
712	11	2.0	335	2	Q59PE4_CANAL	Q59pe4_candida alb	785	11	2.0	426	2	Q961P7_DROME	Q961p7_drosophila
713	11	2.0	335	2	Q06U23_NEUCR	Q06u23_neurospora	786	11	2.0	426	2	Q9EN59_NPVST	Q9en59_spodoptera
714	11	2.0	335	2	Q5510_DICD1	Q5510_dictyosteli	787	11	2.0	427	2	Q559E7_DICD1	Q559e7_dictyosteli
715	11	2.0	335	2	Q06HM2_DICD1	Q06hm2_dictyosteli	788	11	2.0	427	2	Q6ORL2_CAEBR	Q6orl2_caenothabdi
716	11	2.0	341	2	Q2YEO6_EPTST	Q2Yeo6_eptretus	789	11	2.0	427	2	Q06117_DICD1	Q06117_dictyosteli
717	11	2.0	343	2	Q5TQ54_ANOGA	Q5tq54_anopheles g	790	11	2.0	431	2	Q4DYL7_TRYCR	Q4dyl7_rypanosoma
718	11	2.0	344	2	Q76741_DICD1	Q76741_dictyosteli	791	11	2.0	431	2	Q54DX0_DICD1	Q54dx0_dictyosteli
719	11	2.0	351	2	Q54Y66_DICD1	Q54y66_dictyosteli	792	11	2.0	433	2	Q9VFD1_DROME	Q9vfd1_drosophila
720	11	2.0	353	2	Q5A9X0_CANAL	Q5a9x0_candida alb	793	11	2.0	435	2	Q59TR9_CANAL	Q59tr9_candida alb
721	11	2.0	354	2	Q5CRM5_CRYPV	Q5crw5_cryptospori	794	11	2.0	435	2	Q5FMW6_XENLA	Q5fmw6_xenopus lae
722	11	2.0	358	2	Q2YEB01_EPTST	Q2Yeb01_eptretus	795	11	2.0	438	2	Q54SX2_DICD1	Q54sx2_dictyosteli
723	11	2.0	358	2	Q54DH9_DICD1	Q54dh9_dictyosteli	796	11	2.0	438	2	Q552M6_DICD1	Q552m6_dictyosteli
724	11	2.0	362	2	Q7YU06_9TRYP	Q7Yug6_rypanosoma	797	11	2.0	438	2	Q06J06_DICD1	Q06j06_dictyosteli
725	11	2.0	364	2	Q7S2P4_NEUCR	Q7s2p4_neurospora	798	11	2.0	442	2	Q9BY67_HUMAN	Q9by67_homo sapien
726	11	2.0	364	2	Q54NB6_DICD1	Q54nb6_dictyosteli	799	11	2.0	443	2	Q54VW8_DICD1	Q54vw8_dictyosteli
727	11	2.0	364	2	Q7YU05_9TRYP	Q7Yug5_rypanosoma	800	11	2.0	445	2	Q5CV35_CRYPV	Q5cv35_cryptospori
728	11	2.0	364	2	Q068R2_ORYSA	Q068r2_oryza sativ	801	11	2.0	445	2	Q084L1_MOUSE	Q084l1_mus musculu
729	11	2.0	365	2	Q7YU06_9TRYP	Q7Yug6_rypanosoma	802	11	2.0	448	2	Q5CN85_CRYHO	Q5cn85_cryptospori
730	11	2.0	365	2	Q069R5_DICD1	Q069r5_dictyosteli	803	11	2.0	448	2	Q60SF2_CAEBR	Q60sf2_caenothabdi
731	11	2.0	366	2	Q54QC3_DICD1	Q54qc3_dictyosteli	804	11	2.0	448	2	Q0N310_CAEBL	Q0n310_caenothabdi
732	11	2.0	366	2	Q55CN7_DICD1	Q55cn7_dictyosteli	805	11	2.0	452	2	Q54DVO_DICD1	Q54dvo_dictyosteli
733	11	2.0	366	2	Q09YZ40_SIVCZ	Q09yz40_chimpanzee	806	11	2.0	453	2	Q07OS8_NEUCR	Q07os8_neurospora
734	11	2.0	367	2	Q7YU08_9TRYP	Q7Yug8_rypanosoma	807	11	2.0	454	2	Q4W9V5_ASFPV	Q4w9v5_aspergillus
735	11	2.0	369	2	Q54B89_DICD1	Q54b89_dictyosteli	808	11	2.0	454	2	Q54YU7_DICD1	Q54y17_dictyosteli
736	11	2.0	369	2	Q7YU01_9TRYP	Q7Yug1_rypanosoma	809	11	2.0	455	2	Q54H69_DICD1	Q54h69_dictyosteli
737	11	2.0	369	2	Q7YUQ2_9TRYP	Q7Yug2_rypanosoma	810	11	2.0	456	2	Q5CXK6_CRYPV	Q5cxk6_cryptospori
738	11	2.0	369	2	Q7YUQ3_9TRYP	Q7Yug3_rypanosoma	811	11	2.0	456	2	Q0R5M8_MOUSE	Q0r5m8_mus musculu
739	11	2.0	369	2	Q7YUQ4_9TRYP	Q7Yug4_rypanosoma	812	11	2.0	461	2	Q5CF54_CRYHO	Q5cf54_cryptospori
740	11	2.0	371	2	Q59QU7_CANAL	Q59qu7_candida alb	813	11	2.0	463	2	Q54BV8_DICD1	Q54bv8_dictyosteli
741	11	2.0	372	2	Q55SK4_DICD1	Q55sk4_dictyosteli	814	11	2.0	464	2	Q55EB3_DICD1	Q55eb3_dictyosteli
742	11	2.0	373	2	Q59QSS_CANAL	Q59qss_candida alb	815	11	2.0	465	2	Q06KH4_DICD1	Q06kh4_dictyosteli
743	11	2.0	374	2	Q54ZP7_DICD1	Q54zp7_dictyosteli	816	11	2.0	467	1	GSXJH_DICD1	PS1136_dictyosteli
744	11	2.0	375	2	Q55FF6_DICD1	Q55ff6_dictyosteli	817	11	2.0	467	2	Q55A28_DICD1	Q55a28_dictyosteli
745	11	2.0	378	2	Q54W11_DICD1	Q54w11_dictyosteli	818	11	2.0	467	2	Q50HU5_MAIZE	Q50hu5_zea mays (m
746	11	2.0	379	2	Q5CTA6_CRYPV	Q5cta6_cryptospori	819	11	2.0	468	2	Q55279_ADEGX	Q55279_avian adeno
747	11	2.0	383	2	Q61GM1_CAEBR	Q61gm1_caenothabdi	820	11	2.0	469	2	Q4X1K5_ASFPV	Q4x1k5_aspergillus
748	11	2.0	384	2	Q54FA7_DICD1	Q54fa7_dictyosteli	821	11	2.0	470	2	Q60NF4_CAEBR	Q60nf4_caenothabdi
749	11	2.0	385	2	Q7LZB8_ICTPU	Q7lzb8_icalulus p	822	11	2.0	474	2	Q7S290_NEUCR	Q7s290_neurospora
750	11	2.0	386	2	Q01759_PNECA	Q01759_pneumocysti	823	11	2.0	476	2	Q6AYP5_RAT	Q6ayp5_rattus norv
751	11	2.0	386	2	Q0GZB9_CAEBL	Q0gzb9_caenothabdi	824	11	2.0	477	2	Q23993_DROME	Q23993_drosophila
752	11	2.0	387	2	Q54BT1_DICD1	Q54bt1_dictyosteli	825	11	2.0	479	2	Q61AU9_CAEBR	Q61au9_caenothabdi
753	11	2.0	387	2	Q54HD0_DICD1	Q54hd0_dictyosteli	826	11	2.0	481	2	Q557H0_DICD1	Q557h0_dictyosteli
754	11	2.0	390	2	Q0BMP08_BOMMO	Q0bmp08_bomblyx mori	827	11	2.0	482	2	Q7SE49_NEUCR	Q7se49_neurospora
755	11	2.0	391	2	Q54M74_DICD1	Q54m74_dictyosteli	828	11	2.0	484	1	OAR2_LOCMI	Q25322_locusta mlg
756	11	2.0	391	2	Q54XP0_DICD1	Q54xf0_dictyosteli	829	11	2.0	484	1	Q54EB6_DICD1	Q54eb6_dictyosteli
757	11	2.0	392	2	Q0B1C1_PLAUF7	Q0b1c1_plasmodium	830	11	2.0	487	2	Q5AAN3_CANAL	Q5aan3_candida alb
758	11	2.0	392	2	Q069Z58_MOUSE	Q069z58_mus musculu	831	11	2.0	490	2	Q7S3X3_NEUCR	Q7s3x3_neurospora
759	11	2.0	394	2	Q54INS_DICD1	Q54ins_dictyosteli	832	11	2.0	493	2	Q59TN7_CANAL	Q59tn7_candida alb
760	11	2.0	396	2	Q54G38_DICD1	Q54g38_dictyosteli	833	11	2.0	495	2	Q623K5_CAEBR	Q623k5_caenothabdi
761	11	2.0	398	2	Q22902_CAEBL	Q22902_caenothabdi	834	11	2.0	502	2		

835	11	2.0	503	1	NAS14_CABEL	019269 caenorhabdi	908	11	2.0	569	2	04WU5 ASPFU	04w1h5 aspergillus
836	11	2.0	504	2	Q4IB23 GIBZE	04ib23 gibberella	909	11	2.0	570	2	06IN12_CABR	06in12 caenorhabdi
837	11	2.0	504	2	06CCD0 YARLI	06ccd0 yarrowia 1i	910	11	2.0	571	2	05A181_CANAL	05a181 candida alb
838	11	2.0	505	2	04FY06 LEIMA	04fy06 leishmania	911	11	2.0	571	2	08NU09_CANAL	08nu09 candida alb
839	11	2.0	506	2	03UGJ7 MOUSE	03ugj7 mus musculus	912	11	2.0	572	2	05CM19_CRYHO	05cm19 cryptospori
840	11	2.0	507	2	04W907 ASPFU	04w907 aspergillus	913	11	2.0	572	2	04OGJ6_LEIMA	04ogj6 leishmania
841	11	2.0	507	2	05ABR3_DICDI	05abr3 dictyostei	914	11	2.0	573	2	054X06_DICDI	054x06 dictyostei
842	11	2.0	507	2	05CFP0_CRYHO	05cfp0 cryptospori	915	11	2.0	573	2	08TOR9_DROME	08tor9 drosoophila
843	11	2.0	508	2	054L83_DICDI	054l83 dictyostei	916	11	2.0	574	2	05CME1_CRYHO	05cme1 cryptospori
844	11	2.0	510	2	04UAE7_THERN	04uae7 thelletia a	917	11	2.0	574	2	054VD8_DICDI	054vd8 dictyostei
845	11	2.0	511	2	0869L8_DICDI	0869l8 dictyostei	918	11	2.0	574	2	09D5X5_MOUSE	09d5x5 m adult mai
846	11	2.0	512	1	WRK33_ARATH	088p5 arabidopsis	919	11	2.0	576	2	05CVT9_CRYPV	05cvt9 cryptospori
847	11	2.0	512	1	054C79_DICDI	054c79 dictyostei	920	11	2.0	576	2	06IMD1_CABR	06imd1 caenorhabdi
848	11	2.0	513	2	059DV8_DROME	059dv8 drosoophila	921	11	2.0	576	2	054UG9_DICDI	054ug9 dictyostei
849	11	2.0	513	2	06V4A2_DROYA	06v4a2 drosoophila	922	11	2.0	578	2	054S85_DICDI	054s85 dictyostei
850	11	2.0	513	2	086P22_DROME	086p22 drosoophila	923	11	2.0	579	2	07SDE0_NEICR	07sde0 neorospira
851	11	2.0	514	2	09W3K2_DROME	09w3k2 anopheles g	924	11	2.0	580	2	054UB4_DICDI	054ub4 dictyostei
852	11	2.0	514	2	07PDW3_ANOGA	07pdw3 anopheles g	925	11	2.0	580	2	05ZDM0_ORYSA	05zdm0 oryza sativ
853	11	2.0	515	2	059MM9_CANAL	059mm9 candida alb	926	11	2.0	583	2	055EC4_DICDI	055ec4 dictyostei
854	11	2.0	516	2	054ZP8_DICDI	054zp8 dictyostei	927	11	2.0	584	2	05ANT0_DICDI	05ant0 dictyostei
855	11	2.0	517	1	LAIC_DIACA	054f6 dianthus ca	928	11	2.0	585	2	054WV2_DICDI	054wv2 dictyostei
856	11	2.0	518	2	05SD66_DICDI	05sd66 dictyostei	929	11	2.0	585	2	054IK3_DICDI	054ik3 dictyostei
857	11	2.0	518	2	043753_DIACA	043753 dianthus ca	930	11	2.0	588	2	054VP4_DICDI	054vp4 dictyostei
858	11	2.0	519	2	009592_CABEL	009592 caenorhabdi	931	11	2.0	592	2	055GD1_DICDI	055gd1 dictyostei
859	11	2.0	520	2	0871E3_NEICR	0871e3 neorospira	932	11	2.0	594	2	054FP9_DICDI	054fp9 dictyostei
860	11	2.0	520	2	09VHU5_DROME	09vhu5 drosoophila	933	11	2.0	594	2	05CP09_CRYHO	05cp09 cryptospori
861	11	2.0	521	2	054UB1_DICDI	054ub1 dictyostei	934	11	2.0	594	2	05CXZ3_CRYPV	05cxz3 cryptospori
862	11	2.0	522	2	075UX8_DICDI	075ux8 dictyostei	935	11	2.0	595	2	054HF5_DICDI	054hf5 dictyostei
863	11	2.0	524	2	0653P4_ORYSA	0653f4 oryza sativ	936	11	2.0	600	2	017490_ANOGA	017490 anopheles g
864	11	2.0	525	1	NUP62_RAT	0653f4 oryza sativ	937	11	2.0	601	2	054WU6_DICDI	054wu6 dictyostei
865	11	2.0	526	1	NUP62_MOUSE	063650 mus musculu	938	11	2.0	602	2	086J06_DICDI	086j06 dictyostei
866	11	2.0	526	2	05OHU6_WHEAT	05ohu6 triticum ae	939	11	2.0	604	2	054P93_DICDI	054p93 dictyostei
867	11	2.0	526	2	05OHV1_HORVU	05ohv1 hordeum vul	940	11	2.0	612	2	054MS1_DICDI	054ms1 dictyostei
868	11	2.0	526	2	03TFE2_MOUSE	03tfre2 mus musculu	941	11	2.0	618	2	054RS7_DICDI	054rs7 dictyostei
869	11	2.0	526	2	03THR9_MOUSE	03thr9 mus musculu	942	11	2.0	620	2	04D1E0_TRYCR	04d1e0 trypanosoma
870	11	2.0	526	2	03UZ21_MOUSE	03uz21 mus musculu	943	11	2.0	621	2	061AV0_CABR	061av0 caenorhabdi
871	11	2.0	526	2	03UDG5_MOUSE	03udg5 mus musculu	944	11	2.0	624	2	04CMM2_TRYCR	04cmm2 trypanosoma
872	11	2.0	526	2	05FMA9_MOUSE	05f4j9 m nucleopor	945	11	2.0	626	2	054CY5_DICDI	054cy5 dictyostei
873	11	2.0	526	2	08VAG9_WBSV	08vag9 white spot	946	11	2.0	627	2	06ON14_LEICH	06on14 leishmania
874	11	2.0	527	2	054D61_DICDI	054d61 dictyostei	947	11	2.0	628	2	054MC3_DICDI	054mc3 dictyostei
875	11	2.0	527	2	054ND4_DICDI	054nd4 dictyostei	948	11	2.0	629	2	03WP84_CANAL	03wp84 candida alb
876	11	2.0	527	2	054V97_DICDI	054v97 dictyostei	949	11	2.0	632	2	0591Z5_CANAL	0591z5 candida alb
877	11	2.0	527	2	054V00_DICDI	054v00 dictyostei	950	11	2.0	632	2	04QCG0_LEIMA	04qcg0 leishmania
878	11	2.0	527	2	080TAF_WBSV	08qtas white spot	951	11	2.0	630	2	055D16_DICDI	055d16 dictyostei
879	11	2.0	530	2	061209_CABEL	061209 caenorhabdi	952	11	2.0	635	2	054W99_DICDI	054w99 dictyostei
880	11	2.0	530	2	0553G3_DICDI	0553g3 dictyostei	953	11	2.0	636	2	054XZ6_DICDI	054xz6 dictyostei
881	11	2.0	534	2	07KTI5_DROME	07kti5 drosoophila	954	11	2.0	635	2	08WP77_DICDI	08wp77 dictyostei
882	11	2.0	536	2	054TM6_DICDI	054tm6 dictyostei	955	11	2.0	637	2	054XCG_DICDI	054xcg dictyostei
883	11	2.0	537	2	054MJ3_DICDI	054mj3 dictyostei	956	11	2.0	639	2	05B0H8_EBMENT	05b0h8 aspergillus
884	11	2.0	538	2	0751S0_DICDI	0751s0 dictyostei	957	11	2.0	641	2	054MD1_DICDI	054md1 dictyostei
885	11	2.0	539	2	054GL8_DICDI	054gl8 dictyostei	958	11	2.0	647	2	054KAB_DICDI	054kab dictyostei
886	11	2.0	540	2	054WG1_DICDI	054wg1 dictyostei	959	11	2.0	648	2	095OX0_CABEL	095ox0 caenorhabdi
887	11	2.0	542	2	096F00_DICDI	096f00 dictyostei	960	11	2.0	648	2	04QGI2_LEIMA	04qgi2 leishmania
888	11	2.0	542	2	096F00_DICDI	096f00 dictyostei	961	11	2.0	650	1	ANG15 ASPFU	04180 aspergillus
889	11	2.0	543	2	084UZ6_MESCR	084uz6 mesembryant	962	11	2.0	652	2	054S30_DICDI	054s30 dictyostei
890	11	2.0	545	2	059Y97_CANAL	059y97 candida alb	963	11	2.0	656	2	059ZP8_CANAL	059zp8 candida alb
891	11	2.0	547	2	0559P7_DICDI	0559p7 dictyostei	964	11	2.0	656	2	054UP2_DICDI	054up2 dictyostei
892	11	2.0	548	2	075905_NEICR	075905 neorospira	965	11	2.0	657	2	059ZM1_CANAL	059zm1 candida alb
893	11	2.0	549	2	054J12_DICDI	054j12 dictyostei	966	11	2.0	657	2	05MAT1_CANAL	05mat1 candida alb
894	11	2.0	551	1	PPBJ_RAT	PS1740 rattus norv	967	11	2.0	657	2	086J19_DICDI	086j19 dictyostei
895	11	2.0	552	2	019659_CABEL	019659 caenorhabdi	968	11	2.0	658	2	05VB47_HELAN	05vb47 helianthus
896	11	2.0	555	2	054HM2_DICDI	054hm2 dictyostei	969	11	2.0	660	2	094485_DICDI	094485 dictyostei
897	11	2.0	556	2	054WV4_DICDI	054wv4 dictyostei	970	11	2.0	663	2	054DG3_DICDI	054dg3 dictyostei
898	11	2.0	558	2	054TVA_DICDI	054tva dictyostei	971	11	2.0	662	2	03HVK1_9EUKA	03hvk1 monosiga ov
899	11	2.0	559	1	PEBI_MOUSE	P24822 mus musculu	972	11	2.0	667	2	07YYX0_CRYPV	07yyx0 cryptospori
900	11	2.0	559	2	054DK7_DICDI	054dk7 dictyostei	973	11	2.0	668	2	055F29_DICDI	055f29 dictyostei
901	11	2.0	559	2	09VN36_DROME	09vn36 drosoophila	974	11	2.0	668	2	04QGR4_LEIMA	04qgr4 leishmania
902	11	2.0	561	2	059MZ6_DROME	059mz6 candida alb	975	11	2.0	669	2	05ELV3_DROYA	05elv3 drosoophila
903	11	2.0	562	2	09YKS8_RAT	09yks8 rattus norv	976	11	2.0	670	2	054KD6_DICDI	054kd6 dictyostei
904	11	2.0	565	2	05AB58_CANAL	05ab58 candida alb	977	11	2.0	671	2	041L13_GIBZE	041l13 gibberella
905	11	2.0	565	2	054TY0_DICDI	054ty0 dictyostei	978	11	2.0	671	2	054CU9_DICDI	054cu9 dictyostei
906	11	2.0	566	2	05CTF0_CRYPV	05ctf0 cryptospori	979	11	2.0	674	2	054A00_DICDI	054a00 dictyostei
907	11	2.0	568	2	055E28_DICDI	055e28 dictyostei	980	11	2.0	676	2	060YF4_CABER	060yf4 caenorhabdi

```

981 11 2.0 678 2 054K18_DICDI 054K18_dicystostei
982 11 2.0 680 2 093374_CABEL 093374_caenorhabdi
983 11 2.0 681 2 09XUS9_CABEL 09XUS9_caenorhabdi
984 11 2.0 683 2 054T38_DICDI 054T38_dicystostei
985 11 2.0 684 2 054CR3_DICDI 054CR3_dicystostei
986 11 2.0 684 2 08T848_DICDI 08T848_dicystostei
987 11 2.0 685 2 061M14_CABER 061M14_caenorhabdi
988 11 2.0 687 2 055C68_DICDI 055C68_dicystostei
989 11 2.0 687 2 061P32_CABER 061P32_caenorhabdi
990 11 2.0 689 2 054BN2_DICDI 054BN2_dicystostei
991 11 2.0 689 2 09Y114_HABCO 09Y114_haemochus
992 11 2.0 690 2 054BX1_DICDI 054BX1_dicystostei
993 11 2.0 691 2 054LX1_DICDI 054LX1_dicystostei
994 11 2.0 693 2 007241_PNECA 007241_pneumocysti
995 11 2.0 697 1 CRAC_DICDI P35401_pneumocysti
996 11 2.0 698 2 054N17_DICDI 054N17_dicystostei
997 11 2.0 698 2 055BH6_DICDI 055BH6_dicystostei
998 11 2.0 699 2 0552K5_DICDI 0552K5_dicystostei
999 11 2.0 701 2 054XW5_DICDI 054XW5_dicystostei
1000 11 2.0 701 2 09U233_CABEL 09U233_caenorhabdi

```

ALIGNMENTS

RESULT 1

```

09R396_ECOLI PRELIMINARY; PRT; 558 AA.
ID 09R396_ECOLI
AC 09R396;
DT 01-MAY-2000, integrated into UniProtKB/TrEMBL.
DT 01-MAY-2000, sequence version 1.
DT 07-FEB-2006, entry version 25.
DE Translocated intimin receptor Tlr (L0027).
CN Name: tlr;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=86/24;
RC MEDLINE=99242825; PubMed=10225900;
RA Devaney R., Stein M., Reinscheid D., Abe A., Rusechowski S.,
RA Finlay B.B.;
RT "Enterohemorrhagic Escherichia coli O157:H7 produces Tlr, which is
RT translocated to the host cell membrane but is not tyrosine
RT phosphorylated.";
RL Infect. Immun. 67:2389-2398(1999).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC43895, and BDL933;
RC MEDLINE=98339885; PubMed=9673266;
RA Perna N.T., Mayhew G.F., Postel G., Elliott S., Donnenberg M.S.,
RA Kaper J.B., Blatterer F.R.;
RT "Molecular evolution of a pathogenicity island from enterohemorrhagic
RT Escherichia coli O157:H7.";
RL Infect. Immun. 66:3810-3817(1998).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CPG7;
RC PubMed=16272509; DOI=10.1128/JCM.43.11.5715-5720.2005;
RA Garmendia J., Ren Z., Tennant S., Midollì Vierra M.A., Chong Y.,
RA Whaley A., Azopardo K., Dahan S., Stricli M.P., Franzolin M.R.,
RA Trabelsi L.R., Phillips A., Gomes T.A., Xu J., Robins-Browne R.,
RA Frankel G.;
RT "Distribution of tccp in Clinical Enterohemorrhagic and
RT Enteropathogenic Escherichia coli Isolates.";
RL J. Clin. Microbiol. 43:5715-5720(2005).
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
DR EMBL, AF125993; AAD29391.1; -, Genomic_DNA.

```

```

DR EMBL, AF071034; AAC31506.1; -, Genomic_DNA.
DR EMBL, DQ007021; AAY25392.1; -, Genomic_DNA.
DR PIR, A98199; A98199.
DR PIR, B86045; B86045.
DR HSSP, Q9KWH9; 1F02.
DR SMR, Q9KWH9; 269-333.
DR BiOcyC; ECOL83334-1:EC84561-MONOMER; -.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR003536; Tlr_rcpt.
DR Pfam; PF07489; Tlr_receptor_C; 1.
DR Pfam; PF03549; Tlr_receptor_M; 1.
DR Pfam; PF07490; Tlr_receptor_N; 1.
DR PRINTS; PR01370; TRANSINTIMINR.
DR Receptor.
SQ SEQUENCE 558 AA; 58022 MW; 99C417222D4BAA1 CRC64;

```

```

Query Match 100.0%; Score 558; DB 2; Length 558;
Best local similarity 100.0%; Pred. No. 0;
Matches 558; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MPICNLGNPNVNNISIPAPPLPSQTDGAGRGQLINSTPLGSRALFTPVNSMADSGD 60
DB 1 MPICNLGNPNVNNISIPAPPLPSQTDGAGRGQLINSTPLGSRALFTPVNSMADSGD 60
QY 61 NRASDVPLPVPNRLAASEITLNDGFEVLHDHGPLDTLNQIGSSVFRVETOEDGKHA 120
DB 61 NRASDVPLPVPNRLAASEITLNDGFEVLHDHGPLDTLNQIGSSVFRVETOEDGKHA 120
QY 121 VGRNGVETSVLSDQEVARLQSIDPEGDKVFETGGRGAGHAMVTASDITTEARQIL 180
DB 121 VGRNGVETSVLSDQEVARLQSIDPEGDKVFETGGRGAGHAMVTASDITTEARQIL 180
QY 121 VGRNGVETSVLSDQEVARLQSIDPEGDKVFETGGRGAGHAMVTASDITTEARQIL 180
DB 121 VGRNGVETSVLSDQEVARLQSIDPEGDKVFETGGRGAGHAMVTASDITTEARQIL 180
QY 181 ELLEPKGEGESKGESGVGELRENSGANTTETQSTSSLRSPKMLALGTAT 240
DB 181 ELLEPKGEGESKGESGVGELRENSGANTTETQSTSSLRSPKMLALGTAT 240
QY 181 ELLEPKGEGESKGESGVGELRENSGANTTETQSTSSLRSPKMLALGTAT 240
DB 181 ELLEPKGEGESKGESGVGELRENSGANTTETQSTSSLRSPKMLALGTAT 240
QY 241 GLIGLAATGIYQALALTPPEPSPTTDDPAAASATETATRDQLTEAFQNPDMQKNIDE 300
DB 241 GLIGLAATGIYQALALTPPEPSPTTDDPAAASATETATRDQLTEAFQNPDMQKNIDE 300
QY 241 GLIGLAATGIYQALALTPPEPSPTTDDPAAASATETATRDQLTEAFQNPDMQKNIDE 300
DB 241 GLIGLAATGIYQALALTPPEPSPTTDDPAAASATETATRDQLTEAFQNPDMQKNIDE 300
QY 301 LGNAIPSGVLKDDVVAEIEQAKAAGEBAKQAIENNAQAQKDEQQAQOEELKVS 360
DB 301 LGNAIPSGVLKDDVVAEIEQAKAAGEBAKQAIENNAQAQKDEQQAQOEELKVS 360
QY 301 LGNAIPSGVLKDDVVAEIEQAKAAGEBAKQAIENNAQAQKDEQQAQOEELKVS 360
DB 301 LGNAIPSGVLKDDVVAEIEQAKAAGEBAKQAIENNAQAQKDEQQAQOEELKVS 360
QY 361 AGYGLSGALLIGGGIGAVATLALHKKQPVQOTTTTTTTTSARTEVKNPANTPAQG 420
DB 361 AGYGLSGALLIGGGIGAVATLALHKKQPVQOTTTTTTTTSARTEVKNPANTPAQG 420
QY 421 NVDPGSEDTMESRRSSMASTSTFTDSSIGTVQNPYADVKTSLHDSQVPTSNSNTSVQ 480
DB 421 NVDPGSEDTMESRRSSMASTSTFTDSSIGTVQNPYADVKTSLHDSQVPTSNSNTSVQ 480
QY 421 NVDPGSEDTMESRRSSMASTSTFTDSSIGTVQNPYADVKTSLHDSQVPTSNSNTSVQ 480
DB 421 NVDPGSEDTMESRRSSMASTSTFTDSSIGTVQNPYADVKTSLHDSQVPTSNSNTSVQ 480
QY 481 NMGNTDSVVYSTIOHPEDDTNGARLIGNSAGIQSTYARLALSGGRHDMGGLTGSSN 540
DB 481 NMGNTDSVVYSTIOHPEDDTNGARLIGNSAGIQSTYARLALSGGRHDMGGLTGSSN 540
QY 541 SAVNTSNPPAPGSHRFV 558
DB 541 SAVNTSNPPAPGSHRFV 558

```

RESULT 2

```

QY DB77_ECO57 PRELIMINARY; PRT; 558 AA.
ID QY DB77_ECO57
AC QY DB77_ECO57;
DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.
DT 05-JUL-2004, sequence version 1.
DT 07-FEB-2006, entry version 17.
DE Putative translocated intimin receptor protein (Translocated intimin
DE receptor Tlr).
GN Name: tlr; Ordered locus Names=EC84561, z5112;
OS Escherichia coli O157:H7.

```

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=83334;
 RN
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=0157:H7 / EDL933 / ATCC 700927 / EHEC;
 RX MEDLINE=21074935; PubMed=11206551; DOI=10.1093/dnares/8.1.11;
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
 Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
 Posfai G., Hackett J., Klink S., Bourin A., Siao Y., Miller L.,
 Grodeck E.J., Davis N.W., Lim A., Dimantanta E.T., Potamowski K.,
 Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
 Welch R.A., Blattner F.R.;
 RA "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7",
 RT Nature 409:529-533(2001).
 RL Nature 409:529-533(2001).
 RN
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=0157:H7 / Sakai / RIMD 0509952 / EHEC;
 RX MEDLINE=21156231; PubMed=11258796; DOI=10.1093/dnares/8.1.11;
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
 Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
 Ikida T., Takami H., Honda T., Sasaki K., Ogasawara N., Yasunaga T.,
 Kuhara S., Shiba T., Hatori M., Shinagawa H.;
 RA "Complete genome sequence of enterohaemorrhagic Escherichia coli
 O157:H7 and genomic comparison with a laboratory strain K-12.",
 RT 0157:H7 and genomic comparison with a laboratory strain K-12.",
 RL DNA Res. 8:11-22(2001).
 CC
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NonDerivs License
 CC
 CC EMBL; AE005174; AAG58825.1; -; Genomic DNA.
 CC EMBL; BA000007; BAB37984.1; -; Genomic DNA.
 DR SMR; Q7DB77; 269-333.
 DR GO; GO:0004872; F:receptor activity; IEA.
 DR GO; GO:0007155; P:cell adhesion; IEA.
 DR InterPro; IPR003536; T1r rcpt.
 DR Pfam; PF07489; T1r_receptor_C; 1.
 DR Pfam; PF03549; T1r_receptor_M; 1.
 DR Pfam; PF07490; T1r_receptor_N; 1.
 DR PRINTS; PR01370; TRANSITIMINR.
 DR Complete proteome; Receptor.
 SQ SEQUENCE 558 AA; 58022 MW; 99CA17222D4B4A1 CRC64;
 Query Match 100.0%; Score 558; DB 2; Length 558;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 558; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 MPIGNLGNHPNNNSIPPAPPLPSQTDGAGRGQLINSTGPGSRALFTPVNSMADSGD 60
 Db 1 MPIGNLGNHPNNNSIPPAPPLPSQTDGAGRGQLINSTGPGSRALFTPVNSMADSGD 60
 Oy 61 NRASDVPGLPVNPMLAASEITLNDGFEVLHDHGPLDTLNRQIGSSVFRVETOEDGKHIA 120
 Db 61 NRASDVPGLPVNPMLAASEITLNDGFEVLHDHGPLDTLNRQIGSSVFRVETOEDGKHIA 120
 Oy 121 VQQRNGVETSVLSDEYARLOSIDPEGDKVFFTGGRGAGAHAMTVASDITTEARQRL 180
 Db 121 VQQRNGVETSVLSDEYARLOSIDPEGDKVFFTGGRGAGAHAMTVASDITTEARQRL 180
 Oy 181 ELLEPKGTGESKAGESKGVGELRESNSGAENTTETOTSTSSLSRSDPKMLALGTAT 240
 Db 181 ELLEPKGTGESKAGESKGVGELRESNSGAENTTETOTSTSSLSRSDPKMLALGTAT 240
 Oy 241 GLIGLAATGIVQALATLPEDPSPTTTPDPAASATETATRDQLTKEAFONPDOKYNIDE 300
 Db 241 GLIGLAATGIVQALATLPEDPSPTTTPDPAASATETATRDQLTKEAFONPDOKYNIDE 300
 Oy 301 LGNAITSSVLKDDVANIIEGQAAGEBAKQAIENNAQOKKYDQQAQKQBELKVSSG 360
 Db 301 LGNAITSSVLKDDVANIIEGQAAGEBAKQAIENNAQOKKYDQQAQKQBELKVSSG 360
 Oy 361 AGYGSGLALIGGGIGVAVTAALHRKNQPVETQTTTTTTTTTTSARTVENKPNANTPAOG 420
 Db 361 AGYGSGLALIGGGIGVAVTAALHRKNQPVETQTTTTTTTTTTSARTVENKPNANTPAOG 420

Db 361 AGYGSGLALIGGGIGVAVTAALHRKNQPVETQTTTTTTTTTTSARTVENKPNANTPAOG 420
 Oy 421 NVDTPGSEDTMERSSMASTSTFPDTSIGVONPYADVKSLHDSQVPTNSNTSVQ 480
 Db 421 NVDTPGSEDTMERSSMASTSTFPDTSIGVONPYADVKSLHDSQVPTNSNTSVQ 480
 Oy 481 NMGNSTSVYVSTIQHPRPDITDNGARLLGNPSAGISTYARLLASGLRHDMDGLTGGSN 540
 Db 481 NMGNSTSVYVSTIQHPRPDITDNGARLLGNPSAGISTYARLLASGLRHDMDGLTGGSN 540
 Oy 541 SAVNTSNPNPAPGSHRFV 558
 Db 541 SAVNTSNPNPAPGSHRFV 558
 RESULT 3
 ID 085506_ECOLI PRELIMINARY; PRT; 558 AA.
 AC 085506;
 DT 01-NOV-1998. Integrated into UniProtKB/TrEMBL.
 DT 01-NOV-1998. sequence version 1.
 DT 07-FEB-2006. entry version 18.
 DE Translocated intimin receptor T1r.
 GN Name: t1r;
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=562;
 RN
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=958F2.
 RX MEDLINE=99003184; PubMed=9784578;
 RA Paton A.W., Manning P.A., Woodrow M.C., Paton J.C.;
 RT "Translocated intimin receptors (t1r) of Shiga-toxinogenic Escherichia
 coli isolates belonging to serogroups O26, O111, and O157 react with
 sera from patients with hemolytic-uremic syndrome and exhibit marked
 sequence heterogeneity.",
 RL Infect. Immun. 66:5580-5586(1998).
 CC
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NonDerivs License
 CC
 CC EMBL; AF070067; AAC69314.1; -; Genomic DNA.
 DR HSSP; Q9KMH9; 1F02.
 DR GO; GO:0004872; F:receptor activity; IEA.
 DR GO; GO:0007155; P:cell adhesion; IEA.
 DR InterPro; IPR003536; T1r rcpt.
 DR Pfam; PF07489; T1r_receptor_C; 1.
 DR Pfam; PF03549; T1r_receptor_M; 1.
 DR Pfam; PF07490; T1r_receptor_N; 1.
 DR PRINTS; PR01370; TRANSITIMINR.
 DR Receptor.
 SQ SEQUENCE 558 AA; 58176 MW; CA2CDAC94527C2E CRC64;
 Query Match 60.9%; Score 340; DB 2; Length 558;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 440; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Oy 1 MPIGNLGNHPNNNSIPPAPPLPSQTDGAGRGQLINSTGPGSRALFTPVNSMADSGD 60
 Db 1 MPIGNLGNHPNNNSIPPAPPLPSQTDGAGRGQLINSTGPGSRALFTPVNSMADSGD 60
 Oy 61 NRASDVPGLPVNPMLAASEITLNDGFEVLHDHGPLDTLNRQIGSSVFRVETOEDGKHIA 120
 Db 61 NRASDVPGLPVNPMLAASEITLNDGFEVLHDHGPLDTLNRQIGSSVFRVETOEDGKHIA 120
 Oy 121 VQQRNGVETSVLSDEYARLOSIDPEGDKVFFTGGRGAGAHAMTVASDITTEARQRL 180
 Db 121 VQQRNGVETSVLSDEYARLOSIDPEGDKVFFTGGRGAGAHAMTVASDITTEARQRL 180
 Oy 181 ELLEPKGTGESKAGESKGVGELRESNSGAENTTETOTSTSSLSRSDPKMLALGTAT 240
 Db 181 ELLEPKGTGESKAGESKGVGELRESNSGAENTTETOTSTSSLSRSDPKMLALGTAT 240


```
OY 241 GLIGLATGIYQALATPEPDSPTTDPDAASATETATPDQITKEAFONPNQKVNIDE 300
DB 241 GLIGLATGIYQALATPEPDSPTTDPDAASATETATPDQITKEAFONPNQKVNIDE 300
OY 301 LGNAIPSGVLKDDVVANIEBQAKAGEBAKQOAIENNAQAKCKDEQQAQKEELKVS 360
DB 301 LGNAIPSGVLKDDVVANIEBQAKAGEBAKQOAIENNAQAKCKDEQQAQKEELKVS 360
OY 361 AGYGSGLALIIIGGGI GAVATAAHKKNQPVETTTTTTTTTTSKATYENKPNANTPAQG 420
DB 361 AGYGSGLALIIIGGGI GAVATAAHKKNQPVETTTTTTTTTTSKATYENKPNANTPAQG 420
OY 421 NVDFPGSEDTMESRRSSMAST 441
DB 421 NVDFPGSEDTMESRRSSMAST 441

RESULT 4
O4ZIM4_ECOLI PRELIMINARY; PRT; 558 AA.
AC O4ZIM4_
DT 07-JUN-2005, integrated into UniProtKB/TrEMBL.
DT 07-JUN-2005, sequence version 1.
DE Translocated intimin receptor.
GN Name=tir;
OS Escherichia coli;
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CPG122-G57, and CPG6;
RX PubMed=16272509; DOI=10.1128/JCM.43.11.5715-5720.2005;
RA Gamadia J., Ren Z., Tennant S., Midoll Viera M.A., Chong Y.,
RA Whale A., Azopardo K., Dahan S., Sircili M.P., Franzolin M.R.,
RA Trabulsi L.R., Phillips A., Gomes T.A., Xu J., Robins-Browne R.,
RA Frankel G.;
RT "Distribution of tcep in clinical Enterohemorrhagic and
RT Enteropathogenic Escherichia coli Isolates.";
RL J. Clin. Microbiol. 43:5715-5720(2005).
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs license
CC -----
DR EMBL: DQ007020; AAY25391.1; -; Genomic_DNA.
DR EMBL: DQ007019; AAY25390.1; -; Genomic_DNA.
DR SMR: Q4ZIM4; 271-335.
DR GO: GO:0004872; F:receptor activity; IEA.
DR GO: GO:0007155; P:cell adhesion; IEA.
DR InterPro: IPR003536; Tir_recp.
DR Pfam: PF07489; Tir_receptor_C; 1.
DR Pfam: PF03549; Tir_receptor_M; 1.
DR Pfam: PF07490; Tir_receptor_N; 1.
DR PRINTS: PRO1370; TRANSINTIMINR.
KW Receptor.
SQ SEQUENCE 558 AA; 58008 MW; 69E1ADF5E80AFB10 CRC64;

Query Match 40.0%; Score 223; DB 2; Length 558;
Best Local Similarity 100.0%; Pred. No. 8e-211;
Matches 223; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MPITGLGNPNVNNISIPAPPLPSQTDGAGRGQLINSTGLGSRALFTPVNSMADSGD 60
DB 1 MPITGLGNPNVNNISIPAPPLPSQTDGAGRGQLINSTGLGSRALFTPVNSMADSGD 60
OY 61 NRASVPGLPVNPMLAASEITLNDGFEVLHDHGPLDTLNROIGSSVFRVETOEDGKHIA 120
DB 61 NRASVPGLPVNPMLAASEITLNDGFEVLHDHGPLDTLNROIGSSVFRVETOEDGKHIA 120
OY 121 VGORNGVETSVVLSDOEYARLOSIDPEGKDFVFTGGRGAGHAMVTYASDITTEARORIL 180
DB 121 VGORNGVETSVVLSDOEYARLOSIDPEGKDFVFTGGRGAGHAMVTYASDITTEARORIL 180
```

```
DB 121 VGORNGVETSVVLSDOEYARLOSIDPEGKDFVFTGGRGAGHAMVTYASDITTEARORIL 180
OY 181 ELLEPKGTGESKAGESKGVGELRESNGAENTETOTSTSTS 223
DB 181 ELLEPKGTGESKAGESKGVGELRESNGAENTETOTSTSTS 223

RESULT 5
O58188_ECOLI PRELIMINARY; PRT; 574 AA.
AC O58188;
DT 26-APR-2005, integrated into UniProtKB/TrEMBL.
DT 26-APR-2005, sequence version 1.
DE Translocated intimin receptor.
GN Name=tir;
OS Escherichia coli;
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Baismann's Run stream;
RX PubMed=15870341; DOI=10.1128/AEM.71.5.2511-2519.2005;
RA Higgins J.A., Belt K.T., Karns J.S., Russell-Anelli J., Shelton D.R.;
RT "tir- and stx-Positive Escherichia coli in Stream Waters in a
RT Metropolitan Area.";
RL Appl. Environ. Microbiol. 71:2511-2519(2005).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Baismann's Run stream;
RA Hohn C., Karns J.S., Higgins J.A.;
RL Submitted (FEB-2005) to the EMBL/Genbank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs license
CC -----
DR EMBL: AY944737; AAX47730.1; -; Genomic_DNA.
DR SMR: O58188; 287-351.
DR GO: GO:0004872; F:receptor activity; IEA.
DR GO: GO:0007155; P:cell adhesion; IEA.
DR InterPro: IPR003536; Tir_recp.
DR Pfam: PF07489; Tir_receptor_C; 1.
DR Pfam: PF03549; Tir_receptor_M; 1.
DR Pfam: PF07490; Tir_receptor_N; 1.
DR PRINTS: PRO1370; TRANSINTIMINR.
KW Receptor.
SQ SEQUENCE 574 AA; 59314 MW; DC78996D1BA3F5F0 CRC64;

Query Match 35.7%; Score 199; DB 2; Length 574;
Best Local Similarity 100.0%; Pred. No. 4.5e-187;
Matches 199; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MPITGLGNPNVNNISIPAPPLPSQTDGAGRGQLINSTGLGSRALFTPVNSMADSGD 60
DB 1 MPITGLGNPNVNNISIPAPPLPSQTDGAGRGQLINSTGLGSRALFTPVNSMADSGD 60
OY 61 NRASVPGLPVNPMLAASEITLNDGFEVLHDHGPLDTLNROIGSSVFRVETOEDGKHIA 120
DB 61 NRASVPGLPVNPMLAASEITLNDGFEVLHDHGPLDTLNROIGSSVFRVETOEDGKHIA 120
OY 121 VGORNGVETSVVLSDOEYARLOSIDPEGKDFVFTGGRGAGHAMVTYASDITTEARORIL 180
DB 121 VGORNGVETSVVLSDOEYARLOSIDPEGKDFVFTGGRGAGHAMVTYASDITTEARORIL 180
OY 181 ELLEPKGTGESKAGESK 199
DB 181 ELLEPKGTGESKAGESK 199

RESULT 6
O7BHL5_9ENTR PRELIMINARY; PRT; 547 AA.
ID O7BHL5_9ENTR
```



```

AC 07BHL5;
DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.
DT 05-JUL-2004, sequence version 1.
DT 07-FEB-2006, entry version 8.
DE Translocated intimin receptor Tlr.
OS Citrobacter rodentium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Citrobacter.
OX NCBI_TaxId=67825;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=DS100;
RX MEDLINE=21437640; PubMed=11553577;
RA DOI=10.1128/IAI.69.10.6323-6335.2001;
RA Deng W., Li Y., Vallance B.A., Finlay B.B.;
RT "Locus of enterocyte effacement from Citrobacter rodentium: sequence
RT analysis and evidence for horizontal transfer among attaching and
RT infecting pathogens."
RL Infect. Immun. 69:6323-6335 (2001).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs license
CC -----
DR SMR; 07BHL5; 270-334.
DR GO; GO:0004872; P:receptor activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR GO; GO:0004872; P:receptor activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR003536; Tlr_rcpt.
DR Pfam; PF07489; Tlr_receptor_C; 1.
DR Pfam; PF03549; Tlr_receptor_M; 1.
DR Pfam; PF07490; Tlr_receptor_N; 1.
DR PRINTS; PR01370; TRANSINTIMINR.
KW Receptor.
SQ
SEQUENCE 547 AA; 56284 MW; 02CAC6D625FA6EE1 CRC64;

Query Match
Best Local Similarity 100.0%; Score 28; DB 2; Length 547;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 281 DQLTKEAFQNPQKVNIDELGNALPSG 308
Db 282 DQLTKEAFQNPQKVNIDELGNALPSG 309

RESULT 7
Q9ETI1_9ENTR PRELIMINARY; PRT; 547 AA.
AC Q9ETI1;
DT 01-MAR-2001, integrated into UniProtKB/TrEMBL.
DT 01-MAR-2001, sequence version 1.
DT 07-FEB-2006, entry version 12.
DE Translocated intimin receptor Tlr.
OS Citrobacter rodentium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Citrobacter.
OX NCBI_TaxId=67825;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CDC 1843-73T, and DS8100;
RX MEDLINE=20553330; PubMed=1101562;
RA Lupercio S.A., Newman J.V., Dangler C.A., Schrenzel M.D.,
RA Brenner D.J., Steigerwalt A.G., Schauer D.B.;
RT "Citrobacter rodentium, the causative agent of transmissible murine
RT colonic hyperplasia, exhibits clonality: synonymy of C. rodentium and
RT mouse-pathogenic Escherichia coli."
RL J. Clin. Microbiol. 38:4343-4350 (2000).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs license
CC -----
DR EMBL; AF301618; AAG40758.1; -; Genomic DNA.
DR EMBL; AF301617; AAG25642.1; -; Genomic DNA.
DR HSSP; Q9KWH9; 1F02.

```

```

DR SMR; Q9ETI1; 270-334.
DR GO; GO:0004872; P:receptor activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR003536; Tlr_rcpt.
DR Pfam; PF07489; Tlr_receptor_C; 1.
DR Pfam; PF03549; Tlr_receptor_M; 1.
DR Pfam; PF07490; Tlr_receptor_N; 1.
DR PRINTS; PR01370; TRANSINTIMINR.
KW Receptor.
SQ
SEQUENCE 547 AA; 56271 MW; CB831BB301049C37 CRC64;

Query Match
Best Local Similarity 100.0%; Score 28; DB 2; Length 547;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 281 DQLTKEAFQNPQKVNIDELGNALPSG 308
Db 282 DQLTKEAFQNPQKVNIDELGNALPSG 309

RESULT 8
Q9WXK1_ECOLI PRELIMINARY; PRT; 547 AA.
AC Q9WXK1;
DT 01-NOV-1999, integrated into UniProtKB/TrEMBL.
DT 01-NOV-1999, sequence version 1.
DT 07-FEB-2006, entry version 16.
DE Translocated intimin receptor.
GN Name=tlr;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxId=562;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MPEC.
RA Okutani A., Itoh K., Saakawa C.;
RT Submitted (Apr-1999) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs license
CC -----
DR EMBL; AB026719; BAA77400.1; -; Genomic DNA.
DR HSSP; Q9KWH9; 1F02.
DR SMR; Q9WXK1; 270-334.
DR GO; GO:0004872; P:receptor activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR003536; Tlr_rcpt.
DR Pfam; PF07489; Tlr_receptor_C; 1.
DR Pfam; PF03549; Tlr_receptor_M; 1.
DR Pfam; PF07490; Tlr_receptor_N; 1.
DR PRINTS; PR01370; TRANSINTIMINR.
KW Receptor.
SQ
SEQUENCE 547 AA; 56284 MW; 02CAC6D625FA6EE1 CRC64;

Query Match
Best Local Similarity 100.0%; Score 28; DB 2; Length 547;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 281 DQLTKEAFQNPQKVNIDELGNALPSG 308
Db 282 DQLTKEAFQNPQKVNIDELGNALPSG 309

RESULT 9
O85508_ECOLI PRELIMINARY; PRT; 538 AA.
AC O85508;
DT 01-NOV-1998, integrated into UniProtKB/TrEMBL.
DT 01-NOV-1998, sequence version 1.
DT 07-FEB-2006, entry version 16.
DE Translocated intimin receptor Tlr.
GN Name=tlr;

```

```

OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=EPEC87A;
RX MEDLINE=99003184; PubMed=9784578;
RA Paton A.W., Manning P.A., Woodrow M.C., Paton J.C.;
RT "Translocated intimin receptors (Tir) of Shiga-toxinigenic Escherichia
RT coli isolates belonging to serogroups O26, O111, and O157 react with
RT sera from patients with hemolytic-uremic syndrome and exhibit marked
RT sequence heterogeneity.";
RL Infect. Immun. 66:5580-5586(1998).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
DR EMBL; AF070069; AAC69318.1; -; Genomic_DNA.
DR HSSP; Q9KWH9; 1F02.
DR SMR; O85508; 261-325.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR003536; Tir_rcpt.
DR Pfam; PF07489; Tir_receptor_C; 1.
DR Pfam; PF03549; Tir_receptor_M; 1.
DR Pfam; PF07490; Tir_receptor_N; 1.
DR PRINTS; PR01370; TRANSINTIMINR.
KM Receptor.
SQ SEQUENCE 538 AA; 55602 MW; 447052A0E3214D6D CRC64;

Query Match 3.9%; Score 22; DB 2; Length 538;
Best Local Similarity 100.0%; Pred. No. 4.5e-12;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 151 KFVFTGGRGAGHAMVTVASDI 172
Db 152 KFVFTGGRGAGHAMVTVASDI 173

RESULT 10
Q47014_ECOLI PRELIMINARY; PRT; 538 AA.
AC Q47014;
DT 01-NOV-1996, integrated into UniProtKB/TrEMBL.
DT 01-MAY-1999, sequence version 2.
DE Translocated intimin receptor Tir (Translocated intimin co-receptor)
DE (Espe protein).
GN Name=tir; Synonyms=espe;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=REPEC 84/110/1, and B65/56;
RA Krejany E.O.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC MEDLINE=95261;
RX MEDLINE=99003184; PubMed=9784578;
RA Paton A.W., Manning P.A., Woodrow M.C., Paton J.C.;
RT "Translocated intimin receptors (Tir) of Shiga-toxinigenic Escherichia
RT coli isolates belonging to serogroups O26, O111, and O157 react with
RT sera from patients with hemolytic-uremic syndrome and exhibit marked
RT sequence heterogeneity.";
RL Infect. Immun. 66:5580-5586(1998).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=B10;
RX MEDLINE=20187493; PubMed=10722617;

```

```

RX DOI=10.1128/IAI.68.4.2171-2182.2000;
RA Marches O., Nougayrede J.-P., Boullier S., Mainil J., Charlier G.,
RA Raymond I., Pohl P., Boury M., De Rycke J., Milon A., Oswald E.;
RT "Role of Tir and intimin in the virulence of rabbit enteropathogenic
RT Escherichia coli serotype O103:H2.";
RL Infect. Immun. 68:2171-2182(2000).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=413/89-1;
RX MEDLINE=98294040; PubMed=9632251;
RA Deibel C., Kraemer S., Chakraborty T., Ebel F.;
RT "Espe, a novel secreted protein of attaching and effacing bacteria, is
RT directly translocated into infected host cells where it appears as a
RT tyrosine-phosphorylated 90 kDa protein.";
RL Mol. Microbiol. 28:463-474(1998).
RN [5]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=413/89-1;
RA Benkel P., Chakraborty T.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
DR EMBL; U59502; AAC32028.2; -; Genomic_DNA.
DR EMBL; AF070068; AAC69316.1; -; Genomic_DNA.
DR EMBL; AF113597; AAF03080.1; -; Genomic_DNA.
DR EMBL; AJ223063; CA111055.1; -; Genomic_DNA.
DR EMBL; AJ277443; CAC81869.1; -; Genomic_DNA.
DR EMBL; AF132728; AAD27868.1; -; Genomic_DNA.
DR HSSP; Q9KWH9; 1F02.
DR SMR; Q47014; 261-325.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR003536; Tir_rcpt.
DR Pfam; PF07489; Tir_receptor_C; 1.
DR Pfam; PF03549; Tir_receptor_M; 1.
DR Pfam; PF07490; Tir_receptor_N; 1.
DR PRINTS; PR01370; TRANSINTIMINR.
KM Receptor.
SQ SEQUENCE 538 AA; 55421 MW; 31D7A8E227B3D06C CRC64;

Query Match 3.9%; Score 22; DB 2; Length 538;
Best Local Similarity 100.0%; Pred. No. 4.5e-12;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 151 KFVFTGGRGAGHAMVTVASDI 172
Db 152 KFVFTGGRGAGHAMVTVASDI 173

RESULT 11
Q47016_ECOLI PRELIMINARY; PRT; 538 AA.
AC Q47016;
DT 01-NOV-1996, integrated into UniProtKB/TrEMBL.
DT 01-AUG-1999, sequence version 2.
DE Translocated intimin receptor (Tir).
DE Name=tir;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=REPEC 83/39;
RA O'Brien R.A.;
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=REPEC 83/39;
RA Krejany E.O.;

```

RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=RDEC-1;
RC Agin T.S., Boedeker E.C.;
RA Agin T.S., Boedeker E.C.;
RN Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
[4]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=RDEC-1;
RX MEDLINE=98254123; PubMed=9593291;
RA Elliott S.V., Walmsright L.A., McDaniel T.K., Jarvis K.G., Deng Y.K.,
Lai L.C., McNamara B.P., Donenberg M.S., Kaper J.B.;
RT "The complete sequence of the locus of enterocyte effacement (LEE)
from enteropathogenic *Escherichia coli* E2348/69.";
RL Mol. Microbiol. 28:1-4(1998).
[5]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=RDEC-1;
RX MEDLINE=97055784; PubMed=8900070; DOI=10.1016/0378-1097(96)00371-0;
RA Agin T.S., Canney J.R., Boedeker E.C., Wolf M.K.;
RT "Characterization of the *eaeA* gene from rabbit enteropathogenic
Escherichia coli strain RDEC-1 and comparison to other *eaeA* genes from
bacteria that cause attaching-effacing lesions.";
RL FEMS Microbiol. Lett. 144:249-258(1996).
[6]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=RDEC-1;
RX MEDLINE=21153569; PubMed=11254564;
DOI=10.1128/IAI.69.4.2107-2115.2001;
RA Zhu C., Agin T.S., Elliott S.J., Johnson L.A., Thate T.E., Kaper J.B.,
Boedeker E.C.;
RT "Complete nucleotide sequence and analysis of the locus of enterocyte
effacement from rabbit diarrhoeagenic *Escherichia coli* RDEC-1.";
RL Infect. Immun. 69:2107-2115(2001).
[7]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=83/39;
RX MEDLINE=22063667; PubMed=12067342;
RA Tauschek M., Strugnelli R.A., Rodins-Browne R.M.;
RT "Characterization and evidence of mobilization of the LEE
pathogenicity island of rabbit-specific strains of enteropathogenic
Escherichia coli.";
RL Mol. Microbiol. 44:1533-1550(2002).

CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NonDerivs license

DR EMBL, U59504; AAD19750.1; -; Genomic DNA.
DR EMBL, AF045568; AAC15683.1; -; Genomic DNA.
DR EMBL, AF200363; AAK26722.1; -; Genomic DNA.
DR EMBL, AF453441; AAL57549.1; -; Genomic DNA.
DR HSP, O9KM9; 1F02.
DR SMR, O47016; 261-325.
DR GO, GO:0004872; F:receptor activity; IEA.
DR GO, GO:0007155; P:cell adhesion; IEA.
DR InterPro, IPR003536; TIR_rcpt.
DR Pfam, PF07489; TIR_receptor_C; 1.
DR Pfam, PF03549; TIR_receptor_M; 1.
DR Pfam, PF07490; TIR_receptor_N; 1.
DR PRINTS, PRO1370; TRANSINTMINR.
KW Receptor.
SQ SEQUENCE 538 AA, 55411 MW, 3132A969B7B3D06C CRC64;

Query Match 3.9%; Score 22; DB 2; Length 538;
Best Local Similarity 100.0%; Pred. No. 4.5e-12;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 151 KEVFTGRCGAGHAMVTASDI 172
DB 152 KEVFTGRCGAGHAMVTASDI 173

058187 ECOLI
ID Q58187_ECOLI PRELIMINARY; PRT; 538 AA.
AC Q58187;
DT 26-APR-2005, integrated into UniProtKB/TrEMBL.
DT 26-APR-2005, sequence version 1.
DT 07-FEB-2006, entry version 6.
DE Translocated intimin receptor.
GN Name=tir;
OS *Escherichia coli*.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; *Escherichia*.
OX NCBI_TaxId=562;
[1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15870341; DOI=10.1128/AEM.71.5.2511-2519.2005;
RA Higgins J.A., Belt K.T., Karns J.S., Russell-Anelli J., Shelton D.R.;
RT "tir- and *stx*-Positive *Escherichia coli* in Stream Waters in a
Metropolitan Area.";
RL Appl. Environ. Microbiol. 71:2511-2519(2005).
[2]
RP NUCLEOTIDE SEQUENCE.
RA Hohn C., Karns J.S., Higgins J.A.;
RN Submitted (FEB-2005) to the EMBL/GenBank/DBJ databases.

CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NonDerivs license

DR EMBL, AY944738; AAX47731.1; -; Genomic DNA.
DR SMR, Q58187; 261-325.
DR GO, GO:0004872; F:receptor activity; IEA.
DR GO, GO:0007155; P:cell adhesion; IEA.
DR InterPro, IPR003536; TIR_rcpt.
DR Pfam, PF07489; TIR_receptor_C; 1.
DR Pfam, PF03549; TIR_receptor_M; 1.
DR Pfam, PF07490; TIR_receptor_N; 1.
DR PRINTS, PRO1370; TRANSINTMINR.
KW Receptor.
SQ SEQUENCE 538 AA, 55455 MW, D28F5200F04A1890 CRC64;

Query Match 3.9%; Score 22; DB 2; Length 538;
Best Local Similarity 100.0%; Pred. No. 4.5e-12;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 151 KEVFTGRCGAGHAMVTASDI 172
DB 152 KEVFTGRCGAGHAMVTASDI 173

RESULT 13
Q58189_ECOLI PRELIMINARY; PRT; 538 AA.
ID Q58189_ECOLI
AC Q58189;
DT 26-APR-2005, integrated into UniProtKB/TrEMBL.
DT 26-APR-2005, sequence version 1.
DT 07-FEB-2006, entry version 6.
DE Translocated intimin receptor.
GN Name=tir;
OS *Escherichia coli* O157:H-.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; *Escherichia*.
OX NCBI_TaxId=183192;
[1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Gywns Run Gwynbrook stream;
RX PubMed=15870341; DOI=10.1128/AEM.71.5.2511-2519.2005;
RA Higgins J.A., Belt K.T., Karns J.S., Russell-Anelli J., Shelton D.R.;
RT "tir- and *stx*-Positive *Escherichia coli* in Stream Waters in a
Metropolitan Area.";
RL Appl. Environ. Microbiol. 71:2511-2519(2005).
[2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Gywns Run Gwynbrook stream;
RA Hohn C., Shelton D.R., Higgins J.A.;

```
RL Submitted (FEB-2005) to the EMBL/Genbank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs license
CC -----
DR EMBL: AY944736; AAX47729.1; -; Genomic_DNA.
DR SMR: Q58189; 261-325.
DR GO: GO:0004872; F:receptor activity; IEA.
DR GO: GO:0007155; P:cell adhesion; IEA.
DR InterPro: IPR003536; TIR_recp.
DR Pfam: PF07489; TIR_receptor_C; 1.
DR Pfam: PF03549; TIR_receptor_M; 1.
DR Pfam: PF07490; TIR_receptor_N; 1.
DR PRINTS: PR01370; TRANSINTIMINR.
KM Receptor.
SQ SEQUENCE 538 AA; 55496 MW; 9B848C2F508FC943 CRC64;

Query Match 3.9%; Score 22; DB 2; Length 538;
Best Local Similarity 100.0%; Pred. No. 4.5e-12;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 151 KFVFTGGRGAGHAMVTVASDI 172
Db 152 KFVFTGGRGAGHAMVTVASDI 173

RESULT 14
OS B8190_ECOLI PRELIMINARY; PRT; 538 AA.
AC Q58190;
DT 26-APR-2005, integrated into UniProtKB/TrEMBL.
DT 26-APR-2005, sequence version 1.
DT 07-FEB-2006, entry version 6.
DE Translocated intimin receptor.
GN Name=tir;
OS Escherichia coli O157:H-.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=181192;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Spring Branch stream;
RX PubMed=15870341; DOI=10.1128/AEM.71.5.2511-2519.2005;
RA Higgins J.A., Belt K.T., Karns J.S., Russell-Anelli J., Shelton D.R.;
RT "tir- and stx-positive Escherichia coli in Stream Waters in a
RT Metropolitan Area.";
RL Appl. Environ. Microbiol. 71:2511-2519(2005).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Spring Branch stream;
RC Hohn C., Shelton D.R., Higgins J.A.;
RL Submitted (FEB-2005) to the EMBL/Genbank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs license
CC -----
DR EMBL: AY944735; AAX47728.1; -; Genomic_DNA.
DR SMR: Q58190; 261-325.
DR GO: GO:0004872; F:receptor activity; IEA.
DR GO: GO:0007155; P:cell adhesion; IEA.
DR InterPro: IPR003536; TIR_recp.
DR Pfam: PF07489; TIR_receptor_C; 1.
DR Pfam: PF03549; TIR_receptor_M; 1.
DR Pfam: PF07490; TIR_receptor_N; 1.
DR PRINTS: PR01370; TRANSINTIMINR.
KM Receptor.
SQ SEQUENCE 538 AA; 55609 MW; 0336B5E18787C18E CRC64;

Query Match 3.9%; Score 22; DB 2; Length 538;
Best Local Similarity 100.0%; Pred. No. 4.5e-12;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 151 KFVFTGGRGAGHAMVTVASDI 172
```

```
Db 152 KFVFTGGRGAGHAMVTVASDI 173

RESULT 15
OS K5P9_ECOLI PRELIMINARY; PRT; 538 AA.
AC Q5K5P9;
ID Q5K5P9;
DT 15-FEB-2005, integrated into UniProtKB/TrEMBL.
DT 15-FEB-2005, sequence version 1.
DT 07-FEB-2006, entry version 4.
DE Translocated intimin receptor.
GN Name=tir;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=RM1374;
RX MEDLINE=21538660; PubMed=11682182;
RA Jores J., Rumer L., Kieselring S., Kaper J.B., Wieler L.H.;
RT "Identification of a new pathogenicity island inserted in the phev
RT tRNA gene of the bovine Shiga toxin-producing E. coli strain RM1374
RT (O103:H2) harboring a locus of enterocyte effacement that is flanked
RT by intact insertion elements.";
RL Int. J. Med. Microbiol. Lett. 204:75-79(2001).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=RM1374;
RX MEDLINE=22522600; PubMed=12635929;
RA Rumer L., Jores J., Kirsch P., Cavignac Y., Zehmke K., Wieler L.H.;
RT "Dissemination of phev and phev located genomic islands among
RT enteropathogenic (EPEC) and enterohemorrhagic (EHEC) E. coli and their
RT possible role in the horizontal transfer of the locus of enterocyte
RT effacement (LEE).";
RL Int. J. Med. Microbiol. 292:463-475(2003).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=RM1374;
RA Jores J., Wagner S.K., Rumer L., Eichberg J., Laturnus C., Kirsch P.,
RA Scherack P., Teschaepe H., Wieler L.H.;
RT "Description of a 111-kb pathogenicity island (PAI) encoding various
RT virulence features in the enterohemorrhagic E. coli (EHEC) strain
RT RM1374 (O103:H2) and detection of a similar PAI in other EHEC strains
RT of serotype O103:H2.";
RL Int. J. Med. Microbiol. 294:417-425(2005).
RN [4]
RP NUCLEOTIDE SEQUENCE.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs license
CC -----
DR EMBL: AJ303141; CA13867.1; -; Genomic_DNA.
DR SMR: Q5K5P9; 261-325.
DR GO: GO:0004872; F:receptor activity; IEA.
DR GO: GO:0007155; P:cell adhesion; IEA.
DR InterPro: IPR003536; TIR_recp.
DR Pfam: PF07489; TIR_receptor_C; 1.
DR Pfam: PF03549; TIR_receptor_M; 1.
DR Pfam: PF07490; TIR_receptor_N; 1.
DR PRINTS: PR01370; TRANSINTIMINR.
KM Receptor.
SQ SEQUENCE 538 AA; 55482 MW; 7F05D83C6207F117 CRC64;

Query Match 3.9%; Score 22; DB 2; Length 538;
Best Local Similarity 100.0%; Pred. No. 4.5e-12;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 151 KFVFTGGRGAGHAMVTVASDI 172
Db 152 KFVFTGGRGAGHAMVTVASDI 173

Search completed: August 1, 2006, 22:38:30
```

Wed Aug 2 14:25:50 2006

us-09-189-415d-11.01go.rup

Job time : 333 secs

THIS PAGE BLANK (USPTO)